

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2004, 00:50:11 ; Search time 42.7489 Seconds
(without alignments)
3827.903 Million cell updates/sec

Title: US-09-397-967A-17

Perfect score: 25
Sequence: 1 KDYY 4

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/uspto.spool/US09397967/runat.03022004.175631.28576/app.query.fasta_1.1486
-DB=GenEmbl -OPMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=500 -THR_MIN=100 -THR_MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967@cgn.1.1.7598@runat.03022004.175631.28576 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.or:*
21: em.om:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	25	100.0	30	10	MMF10153
2	25	100.0	36	9	S78945
C	3	25	100.0	47	6 AR289401
4	25	100.0	49	10	MUSIGDQBT
C	5	25	100.0	51	6 AX159117
C	6	25	100.0	51	6 AX159118
C	7	25	100.0	60	9 HOMTCRGAA
C	8	25	100.0	63	6 BD035124
C	9	25	100.0	69	3 MIAARN01
C	10	25	100.0	71	6 AR193244
C	11	25	100.0	77	4 AF339962
C	12	25	100.0	77	4 AF339966
C	13	25	100.0	77	4 AF339972
C	14	25	100.0	81	4 AF339971
C	15	25	100.0	83	4 AF339969
C	16	25	100.0	83	4 AF339970
C	17	25	100.0	84	9 AY256110
18	25	100.0	85	9	AY256163
C	19	25	100.0	88	9 HSPA22C6
C	20	25	100.0	89	4 AF339961
C	21	25	100.0	89	4 AF339963
C	22	25	100.0	89	4 AF339964
C	23	25	100.0	89	4 AF339965
C	24	25	100.0	89	4 AF339967
C	25	100.0	89	4	AF339968
C	26	25	100.0	91	8 ATH521235
27	25	100.0	94	6	AX522711
C	28	25	100.0	94	8 ATH521186
C	29	25	100.0	96	9 HSPA22C5
C	30	25	100.0	96	11 HSPA22A11
C	31	25	100.0	97	8 ATH553390
C	32	25	100.0	97	9 HSPA22B5
C	33	25	100.0	99	6 AX571616
C	34	25	100.0	106	11 HSC03P01
35	25	100.0	107	6	I45687
36	25	100.0	107	6	I55779
37	25	100.0	108	6	BD093951
38	25	100.0	108	6	BD178903
C	39	25	100.0	110	6 BD093953
C	40	25	100.0	110	6 BD178905
41	25	100.0	110	9	AF028102
42	25	100.0	117	9	AF028094
C	43	25	100.0	122	6 BD043430
44	25	100.0	123	6	BD045835
45	25	100.0	124	8	ATH525744

ALIGNMENTS

```

MMF10153
LOCUS      30 bp      DNA      linear      ROD 07-MAR-1997
DEFINITION M.musculus immunoglobulin heavy chain CDR3 region (focus 1015 DNA
#3).
ACCESSION  X67378
VERSION     X67378.1 GI:50904
KEYWORDS    diversity region; immunoglobulin; immunoglobulin heavy chain;
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
REFERENCE   1 (bases 1 to 30)
AUTHORS     Jacob, J. and Kelsoe, G.
TITLE       In situ studies of the primary immune response to
JOURNAL     J. Exp. Med. 176 (3), 679-687 (1992)
MEDLINE     92381435
PUBMED      1512536
REFERENCE   2 (bases 1 to 30)
AUTHORS     Jacob, J., Przylepa, J., Miller, C. and Kelsoe, G.
TITLE       In situ studies of the primary immune response to
(4-hydroxy-3-nitrophenyl)acetyl. III. The kinetics of V region
mutation and selection in germinal center B cells
JOURNAL     J. Exp. Med. 178 (4), 1293-1307 (1993)
MEDLINE     93389394
PUBMED      8376935
REFERENCE   3 (bases 1 to 30)
AUTHORS     Jacob, J.
TITLE       Direct Submission
JOURNAL     Submitted (14-JUL-1992) J. Jacob, University of Maryland School of
Medicine, Dept of Microbiology & Immunology, 655 W Baltimore St,
Baltimore MD 21 201, USA
COMMENT     See also X67341-7, X67349-91, J00522, J00529-30, J00532-37 &
J00539.
FEATURES             Location/Qualifiers
     source           1..30
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /cell_line="PAUS-associated focus F1015"
                     /cell_type="B cells"
                     /tissue_type="spleen"
                     /dev_stage="day 10"
                     <1..>30
                     /note="CDR3"
                     /codon_start=1
                     /protein_id="CAA47790.1"
                     /db_xref="GI:1333908"
                     /translation="AKDYYGTGVF"
     CDS             8 a      5 c      9 g      8 t
BASE COUNT      8 a      5 c      9 g      8 t
ORIGIN
Alignment Scores:
Pred. No.:      342      Length:      30
Score:          25.00      Matches:      4
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              10      Gaps:      0
US-09-397-967A-17 (1-4) x MMF10153 (1-30)
QY      1 LysAspTyrTyr 4
Db      4 AGGATTACTAC 15
RESULT 2
LOCUS      36 bp      DNA      linear      PRI 07-MAY-1993
DEFINITION T-cell receptor gamma junctional region Vgamma8/J2.3 [human,
S78945
KEYWORDS    T-cell receptor gamma junctional region Vgamma8/J2.3 [human,
S78945
ACCESSION  S78945
VERSION     S78945.1 GI:244036
KEYWORDS    patient PEER, Genomic, 36 nt].
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 36)
AUTHORS     Breit, T.M., Wolvers-Tettero, I.L., Hahlen, K., van Wering, E.R. and
van Dongen, J.J.
TITLE       Extensive junctional diversity of gamma delta T-cell receptors
expressed by T-cell acute lymphoblastic leukemias: implications for
the detection of minimal residual disease
Leukemia 5 (12), 1076-1086 (1991)
JOURNAL     Leukemia 5 (12), 1076-1086 (1991)
MEDLINE     92130536
PUBMED      1837811
REMARK      GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 78945] from the original journal article.
This sequence comes from Table 3.
FEATURES             Location/Qualifiers
     source           1..36
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            1..36
                     /partial
                     /gene="T-cell receptor gamma junctional region
                     Vgamma8/J2.3"
BASE COUNT      12 a      7 c      7 g      10 t
ORIGIN
Alignment Scores:
Pred. No.:      400      Length:      36
Score:          25.00      Matches:      4
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              9      Gaps:      0
US-09-397-967A-17 (1-4) x S78945 (1-36)
QY      1 LysAspTyrTyr 4
Db      15 AGGATTATTAT 26
RESULT 3
LOCUS      47 bp      DNA      linear      PAT 12-JUN-2003
DEFINITION Sequence 1136 from patent US 6537751.
ACCESSION  AR289401
VERSION     AR289401.1 GI:31676685
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE     1 (bases 1 to 47)
AUTHORS      Cohen, D., Chumakov, I. and Blumenfeld, M.
TITLE        Biallelic markers for use in constructing a high density
disequilibrium map of the human genome
Patent: US 6537751-A 1136 25-MAR-2003;
JOURNAL      Location/Qualifiers
FEATURES     1..47
             /organism="unknown"
BASE COUNT   15 a      7 c      7 g      17 t      1 others
ORIGIN
Alignment Scores:
Pred. No.:      503      Length:      47
Score:          25.00      Matches:      4
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0

```

```

DB:
US-09-397-967A-17 (1-4) x AR289401 (1-47)
QY
  1 LysAspTyrTyr 4
  |||||
  47 AAGGATTACTAT 36
Db

RESULT 4
MUSIGDJBUT
LOCUS
DEFINITION Mus musculus rearranged immunoglobulin heavy chain D-J junction
  (clone 04JH4).
ACCESSION L36768
VERSION L36768.1 GI:927162
KEYWORDS diversity region; immunoglobulin heavy chain; joining region;
  rearranged.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 49)
AUTHORS Chowers, Y., Harwood, J., Holtmeier, W., Kagnoff, M.F. and
  Morzycka-Wroblewska, B.
TITLE No title
JOURNAL Unpublished (1994)
COMMENT Original source text: Mus musculus (clone: 04JH4) bone marrow DNA.
FEATURES
  source
    1..49
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /clone="04JH4"
    /tissue_type="bone marrow"
  misc_feature
    1..49
    /note="D-J junction"
BASE COUNT 13 a 6 c 14 g 16 t
ORIGIN
Alignment Scores:
  Pred. No.: 521 Length: 49
  Score: 25.00 Matches: 4
  Percent Similarity: 100.00% Conservative: 0
  Best Local Similarity: 100.00% Mismatches: 0
  Query Match: 100.00% Indels: 0
  DB: 10 Gaps: 0
US-09-397-967A-17 (1-4) x MUSIGDJBUT (1-49)
QY
  1 LysAspTyrTyr 4
  |||||
  21 AAGGATTACTAT 32
Db

RESULT 5
AX159117/c
LOCUS
DEFINITION Sequence 2445 from Patent WO0140521.
ACCESSION AX159117
VERSION AX159117.1 GI:14540448
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Shimkets, R.A. and Leach, M.
TITLE Nucleic acids containing single nucleotide polymorphisms and
  methods of use thereof
JOURNAL Patent: WO 0140521-A 2445 07-JUN-2001;
  Curagen Corporation (US)
FEATURES
  source
    1..51
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
  misc_feature
    26
    /note="2 of 2 allelic variants (2445 is other entry)"
    Accession number cg39592883"
BASE COUNT 10 a 12 c 9 g 20 t
ORIGIN
Alignment Scores:
  Pred. No.: 539 Length: 51
  Score: 25.00 Matches: 4
  Percent Similarity: 100.00% Conservative: 0
  Best Local Similarity: 100.00% Mismatches: 0
  Query Match: 100.00% Indels: 0
  DB: 6 Gaps: 0
US-09-397-967A-17 (1-4) x AX159117 (1-51)
QY
  1 LysAspTyrTyr 4
  |||||
  44 AAGGATTATTAC 33
Db

RESULT 7
HUMTCRGAA
LOCUS
DEFINITION Human T-cell receptor Ti unproductively rearranged gamma-chain mRNA
  V8-J2-region, partial cds.
ACCESSION M30891
VERSION M30891.1 GI:339401
KEYWORDS J-region; T-cell receptor; V-region; immunoglobulin gamma-chain;
  processed gene.

```

```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
26
/note="1 of 2 allelic variants (2446 is other entry)"
Accession number cg39592883"
BASE COUNT 10 a 11 c 9 g 21 t
ORIGIN
Alignment Scores:
  Pred. No.: 539 Length: 51
  Score: 25.00 Matches: 4
  Percent Similarity: 100.00% Conservative: 0
  Best Local Similarity: 100.00% Mismatches: 0
  Query Match: 100.00% Indels: 0
  DB: 6 Gaps: 0
US-09-397-967A-17 (1-4) x AX159117 (1-51)
QY
  1 LysAspTyrTyr 4
  |||||
  44 AAGGATTATTAC 33
Db

RESULT 6
AX159118/c
LOCUS
DEFINITION Sequence 2446 from Patent WO0140521.
ACCESSION AX159118
VERSION AX159118.1 GI:14540449
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Shimkets, R.A. and Leach, M.
TITLE Nucleic acids containing single nucleotide polymorphisms and
  methods of use thereof
JOURNAL Patent: WO 0140521-A 2446 07-JUN-2001;
  Curagen Corporation (US)
FEATURES
  source
    1..51
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
  misc_feature
    26
    /note="2 of 2 allelic variants (2445 is other entry)"
    Accession number cg39592883"
BASE COUNT 10 a 12 c 9 g 20 t
ORIGIN
Alignment Scores:
  Pred. No.: 539 Length: 51
  Score: 25.00 Matches: 4
  Percent Similarity: 100.00% Conservative: 0
  Best Local Similarity: 100.00% Mismatches: 0
  Query Match: 100.00% Indels: 0
  DB: 6 Gaps: 0
US-09-397-967A-17 (1-4) x AX159117 (1-51)
QY
  1 LysAspTyrTyr 4
  |||||
  44 AAGGATTATTAC 33
Db

RESULT 7
HUMTCRGAA
LOCUS
DEFINITION Human T-cell receptor Ti unproductively rearranged gamma-chain mRNA
  V8-J2-region, partial cds.
ACCESSION M30891
VERSION M30891.1 GI:339401
KEYWORDS J-region; T-cell receptor; V-region; immunoglobulin gamma-chain;
  processed gene.

```

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 60)
AUTHORS    Littman,D.R., Newton,M., Crommie,D., Ang,S.L., Seidman,J.G.,
            Gettner,S.N. and Weiss,A.
TITLE      Characterization of an expressed CD3-associated Ti gamma-chain
JOURNAL    Nature 326 (6108), 85-88 (1987)
MEDLINE    87144613
PUBMED     3102973
COMMENT    Original source text: Human leukemic T-cell line PEER, cDNA to
            mRNA, clone P-gamma-1.
FEATURES   source
            Location/Qualifiers
            1..60
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /map="llq23"
            1..60
            /gene="CD3G"
            <1..>60
            /gene="CD3G"
            /note="TCR Ti gamma-chain"
            /codon_start=1
            /protein_id="AAA61111.1"
            /db_xref="GI:339402"
            /db_xref="GDB:G00-119-765"
            /translation="YYCATWKDYKKLFGSGTTL"
            18..19
            /gene="CD3G"
            /organism="Homo sapiens"
            /organism="Homo sapiens"
            /gene="CD3G"
            /organism="Homo sapiens"
            /organism="Homo sapiens"
            19 a 11 c 12 g 18 t

BASE COUNT 19 a 11 c 12 g 18 t
ORIGIN
Alignment Scores:
Pred. No.: 619 Length: 60
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-397-967a-17 (1-4) x HUMTCRGAA (1-60)

QY 1 LysAspTyrTyr 4
Db 19 AAGGATTATTAT 30

RESULT 8
BD035124/c
LOCUS BD035124 63 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD035124
VERSION BD035124.1 GI:22576866
KEYWORDS JP 2001269182-A/11370.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 63)
AUTHORS Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 11370 02-OCT-2001;
GENSET
COMMENT OS Homo sapiens (human)
PN JP 2001269182-A/11370
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773

```

```

PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCCLAIR, JEAN YVES
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/68 //G06F17/30, C12N15/00, C12N5/00, PC
G06F15/40
CC
FH Key Location/Qualifiers
FEATURES source
            Location/Qualifiers
            1..63
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            18 a 9 c 8 g 28 t

BASE COUNT 18 a 9 c 8 g 28 t
ORIGIN
Alignment Scores:
Pred. No.: 645 Length: 63
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-397-967a-17 (1-4) x BD035124 (1-63)

QY 1 LysAspTyrTyr 4
Db 19 AAGGATTATTAC 8

RESULT 9
MIAARN01/c
LOCUS MIAARN01 69 bp DNA linear INV 11-JUN-2003
DEFINITION Mosquito mitochondrial tRNA-Met.
ACCESSION X00599
VERSION X00599.1 GI:12627
KEYWORDS transfer RNA; transfer RNA-Met.
SOURCE mitochondrial Aedes albopictus (Asian tiger mosquito)
ORGANISM Aedes albopictus
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Aedes.
REFERENCE 1 (bases 1 to 69)
AUTHORS Dubin,D.T. and HsuChen,C.C.
TITLE Sequence and structure of a methionine transfer RNA from mosquito
mitochondria
JOURNAL Nucleic Acids Res. 12 (10), 4185-4189 (1984)
MEDLINE 84221353
COMMENT Data kindly reviewed (15-MAY-1985) by D.T. Dubin.
FEATURES source
            Location/Qualifiers
            1..69
            /organism="Aedes albopictus"
            /organelle="mitochondrion"
            /mol_type="genomic DNA"
            /db_xref="taxon:7160"
            1..69
            /product="tRNA-Met"
            /note="tRNA-Met (omitting terminal CAA)"
            9
            /note="1-methyladenosine"
            /mod_base=mia
            24
            /note="pseudouridine"
            /mod_base=p
            25
            /note="pseudouridine"
            /mod_base=p
            34
            /note="N-(9-beta-D-ribofuranosyl)purine-6-yl)carbamoyl)threonine"
            /mod_base=OTHER
            61
            /note="pseudouridine"

```



```

Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 4                             Gaps: 0

US-09-397-967A-17 (1-4) x AF339966 (1-77)

QY      1 LysAspTyrTyr 4
DB      47 AAAGATTACTAC 36

RESULT 13
AF339972/c
LOCUS   AF339972
DEFINITION Puma concolor clone Pco201 locus FCA035 microsatellite sequence.
ACCESSION AF339972
VERSION   AF339972.1 GI:14348731
KEYWORDS  Puma concolor (puma)
SOURCE    Puma concolor
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Puma.
REFERENCE 1 (bases 1 to 77)
AUTHORS  Culver, M., Menotti-Raymond, M.A. and O'Brien, S.J.
TITLE     Patterns of size homoplasy at 10 microsatellite loci in pumas (Puma concolor)
JOURNAL   Mol. Biol. Evol. 18 (6), 1151-1156 (2001)
MEDLINE   21265424
PUBMED    11371606
REFERENCE 2 (bases 1 to 77)
AUTHORS  Culver, M., Menotti-Raymond, M.A. and O'Brien, S.J.
TITLE     Direct Submission
JOURNAL   Submitted (24-JAN-2001) Fisheries and Wildlife Sciences, VPI&SU, 134 Cheatham, Blacksburg, VA 24061-0321, USA
FEATURES  Location/Qualifiers
            source
            1..77
            /organism="Puma concolor"
            /mol_type="genomic DNA"
            /db_xref="taxon:9696"
            /chromosome="D2"
            /map="locus FCA035"
            /clone="Pco201"
            repeat_region
            1..77
            /note="microsatellite"
            /rpt_type=tandem
            BASE COUNT 19 a 19 g 33 t
            ORIGIN
            1..77
            /organism="Puma concolor"
            /mol_type="genomic DNA"
            /db_xref="taxon:9696"
            /chromosome="D2"
            /map="locus FCA035"
            /clone="Pco201"
            repeat_region
            1..77
            /note="microsatellite"
            /rpt_type=tandem
            BASE COUNT 19 a 19 g 33 t
            ORIGIN

Alignment Scores:
Pred. No.: 766      Length: 77
Score: 25.00        Matches: 4
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 4                             Gaps: 0

US-09-397-967A-17 (1-4) x AF339972 (1-77)

QY      1 LysAspTyrTyr 4
DB      47 AAAGATTACTAC 36

RESULT 14
AF339971/c
LOCUS   AF339971
DEFINITION Puma concolor clone Pco707 locus FCA035 microsatellite sequence.
ACCESSION AF339971
VERSION   AF339971.1 GI:14348730
KEYWORDS  Puma concolor (puma)
SOURCE    Puma concolor
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Puma.

```

```

REFERENCE 1 (bases 1 to 81)
AUTHORS  Culver, M., Menotti-Raymond, M.A. and O'Brien, S.J.
TITLE     Patterns of size homoplasy at 10 microsatellite loci in pumas (Puma concolor)
JOURNAL   Mol. Biol. Evol. 18 (6), 1151-1156 (2001)
MEDLINE   21265424
PUBMED    11371606
REFERENCE 2 (bases 1 to 81)
AUTHORS  Culver, M., Menotti-Raymond, M.A. and O'Brien, S.J.
TITLE     Direct Submission
JOURNAL   Submitted (24-JAN-2001) Fisheries and Wildlife Sciences, VPI&SU, 134 Cheatham, Blacksburg, VA 24061-0321, USA
FEATURES  Location/Qualifiers
            source
            1..81
            /organism="Puma concolor"
            /mol_type="genomic DNA"
            /db_xref="taxon:9696"
            /chromosome="D2"
            /map="locus FCA035"
            /clone="Pco707"
            repeat_region
            1..81
            /note="microsatellite"
            /rpt_type=tandem
            BASE COUNT 19 a 19 g 35 t
            ORIGIN

Alignment Scores:
Pred. No.: 800      Length: 81
Score: 25.00        Matches: 4
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 4                             Gaps: 0

US-09-397-967A-17 (1-4) x AF339971 (1-81)

QY      1 LysAspTyrTyr 4
DB      47 AAAGATTACTAC 36

RESULT 15
AF339969/c
LOCUS   AF339969
DEFINITION Puma concolor clone Pco601 locus FCA035 microsatellite sequence.
ACCESSION AF339969
VERSION   AF339969.1 GI:14348728
KEYWORDS  Puma concolor (puma)
SOURCE    Puma concolor
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Puma.
REFERENCE 1 (bases 1 to 83)
AUTHORS  Culver, M., Menotti-Raymond, M.A. and O'Brien, S.J.
TITLE     Patterns of size homoplasy at 10 microsatellite loci in pumas (Puma concolor)
JOURNAL   Mol. Biol. Evol. 18 (6), 1151-1156 (2001)
MEDLINE   21265424
PUBMED    11371606
REFERENCE 2 (bases 1 to 83)
AUTHORS  Culver, M., Menotti-Raymond, M.A. and O'Brien, S.J.
TITLE     Direct Submission
JOURNAL   Submitted (24-JAN-2001) Fisheries and Wildlife Sciences, VPI&SU, 134 Cheatham, Blacksburg, VA 24061-0321, USA
FEATURES  Location/Qualifiers
            source
            1..83
            /organism="Puma concolor"
            /mol_type="genomic DNA"
            /db_xref="taxon:9696"
            /chromosome="D2"
            /map="locus FCA035"
            /clone="Pco601"
            repeat_region
            1..83
            /note="microsatellite"

```

Thu Feb 5 11:06:36 2004

BASE COUNT 19 a /rpt_type=tandem 22 g 36 t
ORIGIN

Alignment Scores: 817 Length: 83
Pred. No.: 25.00 Matches: 4
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 4

US-09-397-967A-17 (1-4) x AF339969 (1-83)

Qy 1 LysAspTyrTyr 4
| | | | |
Db 47 AAAGATTACTAC 36

Search completed: February 4, 2004, 06:34:16
Job time : 45.7489 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: February 4, 2004, 01:30:36 ; Search time 0.62738 Seconds
(without alignments)
2814.136 Million cell updates/sec

Title: US-09-397-967A-17

Perfect score: 25

Sequence: 1 KDYY 4

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09397967/runat_03022004_175632_28596/app_query.fasta_1.1486
-DB=Issued Patents NA -OFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967@cgn_1.1.81@runat_03022004_175632_28596 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	25	100.0	29	US-09-304-232-625
2	25	100.0	47	Sequence 625, App
3	25	100.0	47	Sequence 1136, Ap
4	25	100.0	71	US-09-422-978-1136
5	25	100.0	107	US-09-363-939A-48
6	25	100.0	107	Sequence 24, Appl
7	25	100.0	107	US-08-441-591-24
8	25	100.0	107	Sequence 24, Appl
9	25	100.0	107	US-08-303-362A-24
10	25	100.0	107	Sequence 41, Appl
11	25	100.0	192	PCT-US95-05600-41
12	25	100.0	192	Sequence 3404, Ap
13	25	100.0	200	US-09-107-532A-3404
14	25	100.0	200	Sequence 146, App
15	25	100.0	200	US-09-495-050A-146
16	25	100.0	270	US-09-328-111-229
17	25	100.0	276	Sequence 744, App
18	25	100.0	276	US-09-134-001C-744
19	25	100.0	282	US-09-313-294A-3240
20	25	100.0	282	Sequence 3240, Ap
21	25	100.0	291	US-09-313-294A-340
22	25	100.0	291	Sequence 340, App

Sequence 2, Appli	301	100.0	3	US-09-388-917-2
Sequence 250, App	301	100.0	4	US-09-439-313-250
Sequence 250, App	301	100.0	4	US-09-352-616A-250
Sequence 250, App	301	100.0	4	US-09-232-149A-250
Sequence 25, Appl	310	100.0	3	US-09-071-035-55
Sequence 29, Appl	311	100.0	3	US-09-284-782-29
Sequence 38, Appl	344	100.0	2	US-08-975-316-38
Sequence 38, Appl	344	100.0	4	US-09-615-192A-38
Sequence 2, Appli	348	100.0	2	US-08-883-344-2
Sequence 18, Appl	348	100.0	4	US-09-872-609-18
Sequence 18, Appl	348	100.0	4	US-09-025-403A-18
Sequence 21, Appl	354	100.0	3	US-08-767-128-21
Sequence 136, App	357	100.0	4	US-09-643-597-196
Sequence 136, App	357	100.0	4	US-09-480-884A-196
Sequence 196, App	357	100.0	4	US-09-542-615A-196
Sequence 196, App	357	100.0	4	US-09-606-421B-196
Sequence 40, Appl	358	100.0	2	US-08-975-316-40
Sequence 40, Appl	358	100.0	4	US-09-615-192A-40
Sequence 54, Appl	366	100.0	3	US-08-908-643C-54
Sequence 19, Appl	372	100.0	4	US-09-672-609-19
Sequence 19, Appl	372	100.0	4	US-09-025-403A-19
Sequence 3, Appli	373	100.0	1	US-08-017-570-3
Sequence 5, Appli	373	100.0	1	US-08-017-570-5
Sequence 3, Appli	373	100.0	1	US-08-471-426-3
Sequence 5, Appli	373	100.0	1	US-08-471-426-5
Sequence 5, Appli	373	100.0	5	PCT-US94-01709-5
Sequence 1119, Ap	375	100.0	4	US-09-107-532A-1119
Sequence 32, Appl	381	100.0	4	US-09-091-725-32
Sequence 1406, Ap	384	100.0	4	US-09-107-532A-1406
Sequence 42, Appl	385	100.0	3	US-08-732-708C-42
Sequence 1230, Ap	390	100.0	4	US-09-107-532A-1230
Sequence 9, Appli	394	100.0	4	US-09-242-913B-9

ALIGNMENTS

RESULT 1

US-09-304-232-625
; Sequence 625, Application US/09304232
; Patent No. 6525185
; GENERAL INFORMATION:
; APPLICANT: Fan, Jian Bing
; APPLICANT: Chakravarti, Aravinda
; APPLICANT: Halushka, Marc Kenneth
; APPLICANT: Case Western Reserve University School of Medicine
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Polymorphisms Associated With
; FILE REFERENCE: 018547-034210US
; CURRENT APPLICATION NUMBER: US/09/304,232
; CURRENT FILING DATE: 1999-05-03
; EARLIER APPLICATION NUMBER: US 60/084,641
; EARLIER FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 909
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 625
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: MREX2 1405
US-09-304-232-625
Alignment Scores:
Pred. No.: 60.2 Length: 29
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-09-397-967A-17 (1-4) x US-09-304-232-625 (1-29)

QY 1 LysAspTyr 4
| | | | |
Db 7 AAGACTATAT 18

RESULT 2

US-09-422-978-1136/c
; Sequence 1136, Application US/09422978
; Patent No. 6537751

GENERAL INFORMATION:

; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET 020CP1
; CURRENT APPLICATION NUMBER: US/09/422,978
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/298,850
; EARLIER FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,732
; EARLIER FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: US 60/082,614
; EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796

SEQ ID NO 1136

LENGTH: 47

TYPE: DNA

ORGANISM: Homo Sapiens

FEATURE:

NAME/KEY: allele

LOCATION: 24

; OTHER INFORMATION: 99-20642-382 : polymorphic base A or G
US-09-422-978-1136

Alignment Scores:

Pred. No.: 96.3 Length: 47
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-397-967A-17 (1-4) x US-09-422-978-1136 (1-47)

QY 1 LysAspTyr 4
| | | | |
Db 47 AAGACTATAT 36

RESULT 3

US-09-363-939A-48/c

; Sequence 48, Application US/09363939A

; Patent No. 6346611

GENERAL INFORMATION:

; APPLICANT: Pagratis, Nikos
; APPLICANT: Lochrie, Michael
; TITLE OF INVENTION: High Affinity TGF β Nucleic Acid Ligands and
; FILE REFERENCE: NEX87
; CURRENT APPLICATION NUMBER: US/09/363,939A
; CURRENT FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 09/046,247
; PRIOR FILING DATE: 1998-03-23
; PRIOR APPLICATION NUMBER: 08/458,424
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: 07/714,131
; PRIOR FILING DATE: 1991-06-10
; PRIOR APPLICATION NUMBER: 07/931,473
; PRIOR FILING DATE: 1992-08-17
; PRIOR APPLICATION NUMBER: 07/964,624
; PRIOR FILING DATE: 1992-10-21
; PRIOR APPLICATION NUMBER: 08/117,991
; PRIOR FILING DATE: 1993-09-08

; PRIOR APPLICATION NUMBER: 07/536,428
; PRIOR FILING DATE: 1990-06-11
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 48

LENGTH: 71

TYPE: RNA

ORGANISM: Artificial Sequence

FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
; NAME/KEY: modified base

LOCATION: (1)..(71)

; OTHER INFORMATION: All pyrimidines are 2'P.

US-09-363-939A-48

Alignment Scores:

Pred. No.: 144 Length: 71
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-397-967A-17 (1-4) x US-09-363-939A-48 (1-71)

QY 1 LysAspTyr 4

| | | | |

Db 26 AAGATTACTAC 15

RESULT 4

US-08-441-591-24

; Sequence 24, Application US/08441591

; Patent No. 5637682

GENERAL INFORMATION:

; APPLICANT: NIEUWLANDT, D., GOLD, L. AND WECKER, M.
; TITLE OF INVENTION: HIGH-AFFINITY
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIGANDS
; TITLE OF INVENTION: TO THE TACHYKININ
; TITLE OF INVENTION: SUBSTANCE P

NUMBER OF SEQUENCES: 66

CORRESPONDENCE ADDRESS:

ADDRESSEE: Swanson & Bratschun, L.L.C.

STREET: 8400 E. Prentice Avenue, Suite 200

CITY: Englewood

STATE: Colorado

COUNTRY: USA

ZIP: 80111

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MG storage

COMPUTER: IBM compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: WordPerfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/441,591

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/303,362

FILING DATE: 9-SEPTEMBER-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/714,131

FILING DATE: 10-JUNE-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/931,473

FILING DATE: 17-AUGUST-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/117,991

FILING DATE: 8-SEPTEMBER 1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/536,428

FILING DATE: 11-JUNE-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX21/C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 107
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-441-591-24

Alignment Scores:
Pred. No.: 214 Length: 107
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-397-967A-17 (1-4) x US-08-441-591-24 (1-107)

Qy 1 LysAspTyrTyr 4
Db 70 AAAGAUUAUAC 81

RESULT 5
US-08-303-362A-24
Sequence 24, Application US/08303362A
Patent No. 5648214
GENERAL INFORMATION:
APPLICANT: NIEUWLANDT, D., GOLD, L. AND WECKER, M.
TITLE OF INVENTION: HIGH-AFFINITY
TITLE OF INVENTION: OLIGONUCLEOTIDE LIGANDS
TITLE OF INVENTION: TO THE TACHYKININ
TITLE OF INVENTION: SUBSTANCE P
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MG storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/303,362A
FILING DATE: 9-SEPTEMBER-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/931,473
FILING DATE: 17-AUGUST-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/117,991
FILING DATE: 8-SEPTEMBER 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992

ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 107
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-303-362A-24

Alignment Scores:
Pred. No.: 214 Length: 107
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-397-967A-17 (1-4) x US-08-303-362A-24 (1-107)

Qy 1 LysAspTyrTyr 4
Db 70 AAAGAUUAUAC 81

RESULT 6
PCT-US95-05600-41
Sequence 41, Application PC/TUS9505600
GENERAL INFORMATION:
APPLICANT: GOLD, LARRY
APPLICANT: NIEUWLANDT, DAN
APPLICANT: WECKER, MATTHEW
APPLICANT: SCHNEIDER, DANIEL J.
APPLICANT: PEIGON, JULI
APPLICANT: ALLEN, PATRICK
APPLICANT: SULLENGER, BRUCE A.
APPLICANT: DOUDNA, JENNIFER, A.
TITLE OF INVENTION: HIGH-AFFINITY LIGANDS OF
TITLE OF INVENTION: INSULIN RECEPTOR ANTIBODIES, TACHYKININ SUBSTANCE
TITLE OF INVENTION: P. HIV INTEGRASE AND HIV-1 REVERSE TRANSCRIPTASE
NUMBER OF SEQUENCES: 239
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MG
MEDIUM TYPE: storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05600
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/238,863
FILING DATE: 06-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/248,632
FILING DATE: 24-MAY-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/303,362
FILING DATE: 09-SEPTEMBER-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/361,795
FILING DATE: 21-DECEMBER-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/117,991
FILING DATE: 08-SEPTEMBER-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/931,473
FILING DATE: 17-AUGUST-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX17/PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3433
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US95-05600-41

Alignment Scores:
Pred. No.: 214 Length: 107
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-397-967A-17 (1-4) x PCT-US95-05600-41 (1-107)

QY 1 LysAspTyrTyr 4
DB 70 AAAGAUUUUAC 81

RESULT 7
US-107-532A-3404
Sequence 3404, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-JUN-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3404:
SEQUENCE CHARACTERISTICS:
LENGTH: 192 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...192
SEQUENCE DESCRIPTION: SEQ ID NO: 3404:
US-09-107-532A-3404

Alignment Scores:
Pred. No.: 379 Length: 192
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-397-967A-17 (1-4) x US-09-107-532A-3404 (1-192)

QY 1 LysAspTyrTyr 4
DB 117 AAAGATTATTAT 128

RESULT 8
US-09-495-050A-146/C
Sequence 146, Application US/09495050A
Patent No. 6492505
GENERAL INFORMATION:
APPLICANT: Roopa, Reddy
APPLICANT: Guegler, Karl, J.
APPLICANT: Au-Young, Janice
TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
FILE REFERENCE: PA-0013 US
CURRENT APPLICATION NUMBER: US/09/495,050A
CURRENT FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/118,318
PRIOR FILING DATE: February 1, 1999
NUMBER OF SEQ ID NOS: 305
SOFTWARE: PERL Program
SEQ ID NO 146
LENGTH: 200
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6492505 1590496CTL
US-09-495-050A-146

Alignment Scores:
Pred. No.: 394 Length: 200
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-397-967A-17 (1-4) x US-09-495-050A-146 (1-200)

Qy 1 LysAspTyrTyr 4

Db 56 AAAGACTATTAT 45

RESULT 9

US-09-328-111-229
; Sequence 229, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 229
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(270)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-229

Alignment Scores:
Pred. No.: 528 Length: 270
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-397-967A-17 (1-4) x US-09-328-111-229 (1-270)

Qy 1 LysAspTyrTyr 4

Db 177 AAAGACTATTAT 188

RESULT 10

US-09-134-001C-744
; Sequence 744, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 744

; LENGTH: 276

; TYPE: DNA

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-744

Alignment Scores:

Pred. No.: 539 Length: 276
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-397-967A-17 (1-4) x US-09-134-001C-744 (1-276)

Qy 1 LysAspTyrTyr 4

Db 141 AAAGACTATTAC 152

RESULT 11

US-09-313-294A-3240
; Sequence 3240, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 3240
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 7006111167H1
; NAME/KEY: unsure
; LOCATION: 29
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-3240

Alignment Scores:
Pred. No.: 550 Length: 282
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-397-967A-17 (1-4) x US-09-313-294A-3240 (1-282)

Qy 1 LysAspTyrTyr 4

Db 11 AAGGATTACTAT 22

RESULT 12

US-09-313-294A-340
; Sequence 340, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A

```

; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 340
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700549028H1
; NAME/KEY: unsure
; LOCATION: 170, 173, 180
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-340

Alignment Scores:
Pred. No.: 567          Length: 291
Score: 25.00           Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4                  Gaps: 0

```

US-09-397-967A-17 (1-4) x US-09-313-294A-340 (1-291)

```

QY 1 LysAspTyrTyr 4
Db 261 AAAGACTACTAT 272

```

RESULT 13

```

US-09-388-917-2/c
; Sequence 2, Application US/09388917
; Patent No. 6258542
; GENERAL INFORMATION:
; APPLICANT: Havashizaki, Yoshihide
; TITLE OF INVENTION: Method for Supporting DNA-Fixation and DNA-Fixed
; FILE REFERENCE: 1794-0120P
; CURRENT APPLICATION NUMBER: US/09/388,917
; CURRENT FILING DATE: 1999-09-02
; EARLIER APPLICATION NUMBER: 10-250619
; EARLIER FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: 11-240910
; EARLIER FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-388-917-2

```

```

Alignment Scores:
Pred. No.: 586          Length: 301
Score: 25.00           Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3                  Gaps: 0

```

US-09-397-967A-17 (1-4) x US-09-388-917-2 (1-301)

```

QY 1 LysAspTyrTyr 4
Db 48 AAAGATTATTAT 37

```

RESULT 14

```

US-09-439-313-250
; Sequence 250, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.

```

```

; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqui
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 250
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-439-313-250

```

```

Alignment Scores:
Pred. No.: 586          Length: 301
Score: 25.00           Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4                  Gaps: 0

```

US-09-397-967A-17 (1-4) x US-09-439-313-250 (1-301)

```

QY 1 LysAspTyrTyr 4
Db 184 AAAGACTACTAT 195

```

RESULT 15

```

US-09-352-616A-250
; Sequence 250, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang, Yuqui
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352,616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 250
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-352-616A-250

```

```

Alignment Scores:
Pred. No.: 586          Length: 301
Score: 25.00           Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4                  Gaps: 0

```

US-09-397-967A-17 (1-4) x US-09-352-616A-250 (1-301)

```

QY 1 LysAspTyrTyr 4
Db 184 AAAGACTACTAT 195

```

Search completed: February 4, 2004, 06:28:05
Job time : 1.62738 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2004, 06:28:12 ; Search time 3.95648 Seconds
(without alignments)
3724.149 Million cell updates/sec

Title: US-09-397-967A-17

Perfect score: 25

Sequence: 1 KDXY 4

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 2449703 seqs, 184181367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFPO.spool/US09397967/runat_03022004_175633_28650/app.query.fasta_1.1486
-DB=Published Applications NA -QFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09397967@cgn2_1.1.383 @runat_03022004_175633_28650
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description

1	25	100.0	25	12	US-10-361-208-442	Sequence 442, App
2	25	100.0	25	15	US-10-215-112-8501	Sequence 8501, App
3	25	100.0	25	15	US-10-098-263B-26149	Sequence 26149, A
4	25	100.0	25	15	US-10-098-263B-35352	Sequence 35352, A
5	25	100.0	25	15	US-10-098-263B-43800	Sequence 43800, A
6	25	100.0	25	15	US-10-098-263B-56890	Sequence 56890, A
7	25	100.0	25	15	US-10-098-263B-63720	Sequence 63720, A
8	25	100.0	25	15	US-10-098-263B-75996	Sequence 75996, A
9	25	100.0	25	15	US-10-098-263B-86117	Sequence 86117, A
10	25	100.0	25	15	US-10-098-263B-95846	Sequence 95846, A
11	25	100.0	25	15	US-10-098-263B-105518	Sequence 105518, A
12	25	100.0	25	15	US-10-098-263B-115432	Sequence 115432, A
13	25	100.0	29	13	US-10-336-638-625	Sequence 625, App
14	25	100.0	30	12	US-10-361-208-63	Sequence 63, Appl
15	25	100.0	35	13	US-10-153-244-263	Sequence 263, App
16	25	100.0	39	12	US-10-361-208-433	Sequence 433, App
17	25	100.0	39	12	US-10-361-208-437	Sequence 437, App
18	25	100.0	39	12	US-10-361-208-441	Sequence 441, App
19	25	100.0	42	12	US-10-361-208-65	Sequence 65, Appl
20	25	100.0	42	12	US-10-361-208-66	Sequence 66, Appl
21	25	100.0	42	12	US-10-361-208-198	Sequence 198, App
22	25	100.0	42	12	US-10-361-208-199	Sequence 199, App
23	25	100.0	47	12	US-10-349-143-1136	Sequence 1136, App
24	25	100.0	99	10	US-09-960-352-13299	Sequence 13299, A
25	25	100.0	109	9	US-09-770-696-802	Sequence 802, App
26	25	100.0	120	13	US-10-027-632-51768	Sequence 51768, A
27	25	100.0	120	14	US-10-027-632-51768	Sequence 51768, A
28	25	100.0	131	10	US-09-969-373-450	Sequence 450, App
29	25	100.0	131	10	US-09-969-373-451	Sequence 451, App
30	25	100.0	138	13	US-10-091-007-163	Sequence 163, App
31	25	100.0	148	10	US-09-969-373-836	Sequence 836, App
32	25	100.0	150	9	US-09-922-217-750	Sequence 750, App
33	25	100.0	150	10	US-09-833-263-750	Sequence 750, App
34	25	100.0	153	14	US-10-025-380-750	Sequence 750, App
35	25	100.0	153	9	US-09-864-761-26750	Sequence 26750, A
36	25	100.0	167	10	US-09-796-692-7365	Sequence 7365, App
37	25	100.0	167	12	US-10-057-475B-7365	Sequence 7365, App
38	25	100.0	167	12	US-10-154-884B-7365	Sequence 7365, App
39	25	100.0	167	15	US-10-040-862-7365	Sequence 7365, App
40	25	100.0	177	12	US-09-864-408A-7645	Sequence 7645, App
41	25	100.0	183	10	US-09-867-701-8586	Sequence 8586, App
42	25	100.0	184	9	US-09-864-761-29922	Sequence 29922, A
43	25	100.0	186	9	US-09-770-696-251	Sequence 251, App
44	25	100.0	186	11	US-09-754-853A-136	Sequence 136, App
45	25	100.0	189	10	US-09-524-035A-619	Sequence 619, App

ALIGNMENTS

RESULT 1
US-10-361-208-442
; Sequence 442, Application US/10361208
; Publication No. US20040009167A1
; GENERAL INFORMATION:
; APPLICANT: Rider, Todd H.
; TITLE OF INVENTION: ANTI-PATHOGEN TREATMENTS
; FILE REFERENCE: 0050.2041-003
; CURRENT APPLICATION NUMBER: US/10/361.208
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,359
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US 60/355,022
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US 60/432,386
; PRIOR FILING DATE: 2002-12-10
; NUMBER OF SEQ ID NOS: 473
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 442
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: primer

```
US-10-361-208-442
Alignment Scores:
Pred. No.: 320
Score: 25.00
Length: 25
Matches: 4
Conservative: 0
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0
DB: 12

US-09-397-967A-17 (1-4) x US-10-361-208-442 (1-25)

QY 1 LysAspTyrTyr 4
Db 14 AAGGACTACTAC 25

RESULT 2
US-10-215-112-8501/c
; Sequence 8501, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Mitmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8501
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-8501
Alignment Scores:
Pred. No.: 320
Score: 25.00
Length: 25
Matches: 4
Conservative: 0
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0
DB: 15

US-09-397-967A-17 (1-4) x US-10-215-112-8501 (1-25)

QY 1 LysAspTyrTyr 4
Db 13 AAGGACTACTAC 2

RESULT 3
US-10-098-263B-26149/c
; Sequence 26149, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 26149
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-26149
Alignment Scores:
Pred. No.: 320
Length: 25
Matches: 4
Conservative: 0
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0
```

```
Score: 25.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0
DB: 15

US-09-397-967A-17 (1-4) x US-10-098-263B-26149 (1-25)

QY 1 LysAspTyrTyr 4
Db 24 AAGGACTACTAC 13

RESULT 4
US-10-098-263B-35352
; Sequence 35352, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 35352
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-35352
Alignment Scores:
Pred. No.: 320
Length: 25
Matches: 4
Score: 25.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0
DB: 15

US-09-397-967A-17 (1-4) x US-10-098-263B-35352 (1-25)

QY 1 LysAspTyrTyr 4
Db 3 AAGGACTACTAT 14

RESULT 5
US-10-098-263B-43800
; Sequence 43800, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 43800
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-43800
Alignment Scores:
Pred. No.: 320
Length: 25
Score: 25.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0
DB: 15
```

```
US-09-397-967A-17 (1-4) x US-10-098-263B-43800 (1-25)

Qy 1 LysAspTyrTyr 4
Db 9 AAGACTACTAC 20

RESULT 6
US-10-098-263B-56890
; Sequence 56890, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 56890
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-56890

Alignment Scores:
Pred. No.: 320 Length: 25
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-397-967A-17 (1-4) x US-10-098-263B-56890 (1-25)

Qy 1 LysAspTyrTyr 4
Db 4 AAGACTACTAC 15

RESULT 7
US-10-098-263B-63720
; Sequence 63720, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 63720
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-63720

Alignment Scores:
Pred. No.: 320 Length: 25
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-397-967A-17 (1-4) x US-10-098-263B-63720 (1-25)

Qy 1 LysAspTyrTyr 4
Db 1 LysAspTyrTyr 4
Db 7 AAGACTACTAC 18
```

```
Db 5 AAGACTACTAC 16

RESULT 8
US-10-098-263B-75996
; Sequence 75996, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 75996
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-75996

Alignment Scores:
Pred. No.: 320 Length: 25
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-397-967A-17 (1-4) x US-10-098-263B-75996 (1-25)

Qy 1 LysAspTyrTyr 4
Db 6 AAGACTACTAT 17

RESULT 9
US-10-098-263B-86117
; Sequence 86117, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 86117
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-86117

Alignment Scores:
Pred. No.: 320 Length: 25
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-397-967A-17 (1-4) x US-10-098-263B-86117 (1-25)

Qy 1 LysAspTyrTyr 4
Db 7 AAGACTACTAT 18

RESULT 10
US-10-098-263B-95846
; Sequence 95846, Application US/10098263B
```

```
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 95846
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-95846

Alignment Scores:
Pred. No.: 320          Length: 25
Score: 25.00           Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15                Gaps: 0

US-09-397-967A-17 (1-4) x US-10-098-263B-95846 (1-25)

QY 1 LysAspTyrTyr 4
Db 11 AAAGACTACTAT 22

RESULT 11
US-10-098-263B-105518
; Sequence 105518, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 105518
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-105518

Alignment Scores:
Pred. No.: 320          Length: 25
Score: 25.00           Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15                Gaps: 0

US-09-397-967A-17 (1-4) x US-10-098-263B-105518 (1-25)

QY 1 LysAspTyrTyr 4
Db 3 AGGACTATTATC 14

RESULT 12
US-10-098-263B-115432/c
; Sequence 115432, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
```

```
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 115432
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-115432

Alignment Scores:
Pred. No.: 320          Length: 25
Score: 25.00           Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15                Gaps: 0

US-09-397-967A-17 (1-4) x US-10-098-263B-115432 (1-25)

QY 1 LysAspTyrTyr 4
Db 13 AAAGACTATTAC 2

RESULT 13
US-10-336-638-625
; Sequence 625, Application US/10336638
; Publication No. US20030170699A1
; GENERAL INFORMATION:
; APPLICANT: Fan, Jian Bing
; APPLICANT: Chakravarti, Aravinda
; APPLICANT: Halushka, Marc Kenneth
; APPLICANT: Case Western Reserve University School of Medicine
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Polymorphisms Associated With
; FILE REFERENCE: 018547-034210US
; CURRENT APPLICATION NUMBER: US/10/336,638
; CURRENT FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: US/09/304,232
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/084,641
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 909
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 625
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: MRLEX2 1405
US-10-336-638-625

Alignment Scores:
Pred. No.: 374          Length: 29
Score: 25.00           Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13                Gaps: 0

US-09-397-967A-17 (1-4) x US-10-336-638-625 (1-29)

QY 1 LysAspTyrTyr 4
Db 7 AAAGACTAYTAT 18

RESULT 14
US-10-361-208-63
; Sequence 63, Application US/10361208
; Publication No. US20040009167A1
```


Search completed: February 4, 2004, 11:29:00
Job time : 3.95648 secs

```

; GENERAL INFORMATION:
; APPLICANT: Rider, Todd H.
; TITLE OF INVENTION: ANTI-PATHOGEN TREATMENTS
; FILE REFERENCE: 0050.2041-003
; CURRENT APPLICATION NUMBER: US/10/361,208
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,359
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US 60/355,022
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US 60/432,386
; PRIOR FILING DATE: 2002-12-10
; NUMBER OF SEQ ID NOS: 473
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: primer
US-10-361-208-63

```

```

Alignment Scores:
Pred. No.: 388 Length: 30
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

```

US-09-397-967A-17 (1-4) x US-10-361-208-63 (1-30)

```

Qy 1 LysAspTyrTyr 4
Db 19 AAAGATTACTAC 30

```

RESULT 15

```

US-10-153-244-263
; Sequence 263, Application US/10153244
; Publication No. US20030144191A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL TRP CHANNEL FAMILY MEMBER, TRP-PL
; FILE REFERENCE: D0144 NP
; CURRENT APPLICATION NUMBER: US/10/153,244
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,599
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/362,944
; PRIOR FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 335
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 263
; LENGTH: 35
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-153-244-263

```

```

Alignment Scores:
Pred. No.: 456 Length: 35
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

```

US-09-397-967A-17 (1-4) x US-10-153-244-263 (1-35)

```

Qy 1 LysAspTyrTyr 4
Db 20 AAAGATTATTAT 31

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2004, 02:45:11 ; Search time 23.4633 Seconds
(without alignments)
4143.406 Million cell updates/sec

Title: US-09-397-967A-17

Perfect score: 25

Sequence: 1 KDYY 4

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ps2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09397967/runat_03022004_175632_28586/app_query.fasta_1.1486
-DB=EST -QFMT=fastap -SUFFIX=ps2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HSPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967 -CGCN 1_1_4514 -runat_03022004_175632_28586 -NCPU=6 -ICPU=3
-NO MAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	25	100.0	28	29	CC459528	SALK 1303
C 2	25	100.0	45	12	BJ028703	BJ028703
C 3	25	100.0	48	14	CD457715	FG04d_030
C 4	25	100.0	48	29	BZ765762	SALK_1342
C 5	25	100.0	50	9	AUI07400	AUI07400
C 6	25	100.0	58	9	AA129203	zn36d03.r
C 7	25	100.0	59	9	AA574548	vm29f09.r
C 8	25	100.0	61	9	AA770896	vt13a11.r
C 9	25	100.0	61	9	AA526231	ni75a02.s
C 10	25	100.0	64	9	AA595043	no40e11.s
C 11	25	100.0	65	10	BG656411	ib37b02.x
C 12	25	100.0	67	9	AA388894	mpl3h03.r
C 13	25	100.0	68	28	BH792714	SALK 0649
C 14	25	100.0	69	29	AL755459	Arabidops
C 15	25	100.0	72	29	AL764243	Arabidops
C 16	25	100.0	73	28	AZ440056	2M0230N22
C 17	25	100.0	75	28	AZ774526	2M0004M05
C 18	25	100.0	76	28	BH849410	SALK 0696
C 19	25	100.0	76	28	BH849415	SALK 0696
C 20	25	100.0	76	29	BX536177	Arabidops
C 21	25	100.0	79	12	B1446457	dai88g10
C 22	25	100.0	80	28	BH850632	SALK 0716
C 23	25	100.0	81	9	AA080299	SWB51A031
C 24	25	100.0	82	10	BF648113	NF04BD09E
C 25	25	100.0	82	28	BH850750	SALK 0717
C 26	25	100.0	83	13	BH834244	T058G01 P
C 27	25	100.0	85	28	AZ767733	1M0567K11
C 28	25	100.0	87	9	A1537909	tp26f01.x
C 29	25	100.0	94	29	AL952445	Arabidops
C 30	25	100.0	97	14	CD406484	Gm CK3135
C 31	25	100.0	98	28	AZ639208	1M0499102
C 32	25	100.0	100	9	AA206256	zg54e12.r
C 33	25	100.0	100	14	CA907810	PCS04037
C 34	25	100.0	101	9	AT000602	AT000602
C 35	25	100.0	101	12	BF1322315	xx19d07.Y
C 36	25	100.0	102	10	BF155669	QV4-HT089
C 37	25	100.0	103	13	BQ785901	saG61a07.
C 38	25	100.0	103	28	AZ922716	SLCot3E07
C 39	25	100.0	104	12	B7430536	B7430536
C 40	25	100.0	104	12	BM874367	laa09g02.
C 41	25	100.0	104	14	CA829282	3529_1.1
C 42	25	100.0	105	10	BE702365	CM4 -NN102
C 43	25	100.0	105	29	CC390859	PUHEV29TD
C 44	25	100.0	106	9	A1974023	sd15d03.Y
C 45	25	100.0	106	10	BE344906	946029H02

ALIGNMENTS

RESULT 1
LOCUS CC459528/c
DEFINITION SALK_130302.51.20.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_130302.51.20.x, genomic survey sequence.
ACCESSION CC459528
VERSION CC459528.1 GI:31225991
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 28)
 REFERENCE
 Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
 C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shim,P.,
 Zimmermann,J. and Ecker,J.R.
 A Sequence-Indexed Library of Insertion Mutations in the
 Arabidopsis Genome

UNPUBLISHED
 CONTACT: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGNAL)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: ecker@salk.edu

Journal
 COMMENT

This is single pass sequence recovered from the left border of
 TDNA.
 Class: TDNA tagged.
 Location/Qualifiers
 1. .28
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone="SALK 130302.51.20.x"
 /clone_lib="Arabidopsis thaliana TDNA insertion lines"
 /notes="PCR was performed on Arabidopsis thaliana lines
 each of which contains one or more TDNA insertion
 elements. The resultant fragment for each line was
 directly sequenced to determine the genomic sequence at
 the site of insertion. Details of the protocols used can
 be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT
 ORIGIN

Alignment Scores:
 Pred. No.: 1.66e+03 Length: 28
 Score: 25.00 Matches: 4
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 29 Gaps: 0

US-09-397-967a-17 (1-4) x CC459528 (1-28)
 QY 1 LysAspTyrTyr 4
 Db 21 AAAGATTATTAT 10

RESULT 2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855

Expressed genes in X. laevis embryo

UNPUBLISHED

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Expressed genes in X. laevis embryo

UNPUBLISHED

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Expressed genes in X. laevis embryo

UNPUBLISHED

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Expressed genes in X. laevis embryo

UNPUBLISHED

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1. .45

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="X1010a10"

/tissue type="whole embryo"

/dev stage="stage 15"

/clone_lib="NIBB Mochii normalized Xenopus neurula

library"

BASE COUNT 16 a 9 c 6 g 9 t 5 others

ORIGIN

Alignment Scores:

Pred. No.: 2.86e+03 Length: 45

Score: 25.00 Matches: 4

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 12 Gaps: 0

US-09-397-967a-17 (1-4) x BJ028703 (1-45)

QY 1 LysAspTyrTyr 4

Db 32 AAGGACTATTAC 43

RESULT 3

CD457715

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Ouellet, Therese

Eastern Cereal and Oilseed Research Centre

Agriculture and Agri-food Canada

Meatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A 0C6,

CANADA

Tel: (613) 759-1658

Fax: (613) 759-1701

Email: ouellet@em.agr.ca.

Location/Qualifiers

1. .48

/organism="Gibberella zeae"

/mol_type="mRNA"

/strain="DAOM 180378"

/db_xref="taxon:5518"

/clone="Fg04d_03o10"

/tissue type="Mycelium"

/dev stage="Asexual"

/lab_host="E. coli"

/clone_lib="Fg04 AAFRC ECORC Fusarium graminearum mycelium

grown on wheat heads"

/notes="Vector: Bluescript SK+; Site 1: EcoRI; Site 2: XhoI

; Fusarium grown on wheat (cv. Roblin) under high

humidity. cDNA made using Stratagene kit."

BASE COUNT 20 a 2 c 9 g 16 t 1 others

ORIGIN

```

Alignment Scores:
Pred. No.: 3.08e+03 Length: 48
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-397-967A-17 (1-4) x CD457715 (1-48)

Qy 1 LysAspTYrTYr 4
Db 2 AAAGATTACTAT 13

RESULT 4
LOCUS BZ765762 48 bp DNA linear GSS 13-MAR-2003
DEFINITION SALK_134270.26.40.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_134270.26.40.x, genomic
survey sequence.
ACCESSION BZ765762
VERSION BZ765762.1 GI:28938315
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 48)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
,Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
JOURNAL Unpublished
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated intron of At2g9130.
Class: TDNA tagged.
FEATURES
source
1..48
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_134270.26.40.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna\_protocols.html"
BASE COUNT 18 a 12 c 6 g 12 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

Alignment Scores:
Pred. No.: 3.08e+03 Length: 48
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 29 Gaps: 0

US-09-397-967A-17 (1-4) x BZ765762 (1-48)

```

```

Qy 1 LysAspTYrTYr 4
Db 33 AAAGATTACTAT 22

RESULT 5
LOCUS AU107400 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU107400 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
COL01196, mRNA sequence.
ACCESSION AU107400
VERSION AU107400.1 GI:13556921
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tanoda,T., Mizushima-Sugano,J., Sese,J., Hata
,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES
source
1..50
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="COL01196"
/clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT 14 a 9 c 7 g 20 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

Alignment Scores:
Pred. No.: 3.22e+03 Length: 50
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-397-967A-17 (1-4) x AU107400 (1-50)

Qy 1 LysAspTYrTYr 4
Db 33 AAAGATTACTAT 44

RESULT 6
LOCUS AA129203 58 bp mRNA linear EST 01-AUG-1997
DEFINITION AA129203 r1 Stratagene endothelial cell 937223 Homo sapiens cDNA
clone IMAGE:549509 5' similar to gb:J03040 SPARC PRECURSOR (HUMAN
), mRNA sequence.
ACCESSION AA129203
VERSION AA129203.1 GI:1689055
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 58)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chissee,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins

```

M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M. Generation and analysis of 280,000 human expressed sequence tags Genome Res. 6 (9), 807-828 (1996)

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1475 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 1.
Location/Qualifiers

FEATURES
source

```
1. .58
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3928026"
/db_xref="taxon:9606"
/clone="IMAGE:549509"
/dev_stage="umbilical vein, 1 passage"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene endothelial cell 937223"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Umbilical vein endothelial cells, passaged once. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3"
```

BASE COUNT 16 a 13 c 19 g 8 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 3.82e+03 Length: 58
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-397-967A-17 (1-4) x AA129203 (1-58)

QY 1 LysAspTyTyr 4
Db 46 AAGGACTATTAC 57

RESULT 7
AA574548/c

LOCUS
DEFINITION VM29f09.r1 Knowles Solter mouse blastocyst B1 Mus musculus cDNA
clone IMAGE:991625 5', mRNA sequence.

ACCESSION AA574548
VERSION AA574548.1 GI:2349174

KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished
COMMENT Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project
Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:563905

Putative full length read
vector to vector length is 101.

FEATURES
source

```
1. .59
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:991625"
/tissue_type="blastocyst"
/dev_stage="embryo (pre-implantation)"
/lab_host="DH10B"
/clone_lib="Knowles Solter mouse blastocyst B1"
/note="Organ: embryo; Vector: pSPORT; Site 1: NotI; Site 2: Sali; Cloned unidirectionally from mRNA prepared from 800 blastocysts. Primer: Sali(dT): 5'-CGTGCACGTCGACGTTTTTTTTTT-3'. cDNAs were cloned into the NotI/Sali sites of a pSPORT vector (Life Technologies). Two different size selections: B1 (larger inserts) and B3."
```

BASE COUNT 21 a 9 c 9 g 20 t
ORIGIN

Alignment Scores:
Pred. No.: 3.9e+03 Length: 59
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-397-967A-17 (1-4) x AA574548 (1-59)

QY 1 LysAspTyTyr 4
Db 33 AAGGACTACTAT 22

RESULT 8
AA770896

LOCUS
DEFINITION vt13a11.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone

IMAGE:1162940 5', similar to SW:KGUA_MOUSE Q64520 GUANYLATE KINASE
,, mRNA sequence.

ACCESSION AA770896
VERSION AA770896.1 GI:2822707

KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished
COMMENT Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1103180"
 /sex="male, pooled"
 /tissue_type="prostate tumor"
 /lab_host="SOLR (kanamycin resistant)"
 /clone_lib="NCI CGAP Pr23"
 /note="Organ: prostate; Vector: Bluescript SK-; Site_1:
 EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
 Oligo dt. Pooled prostate tumors. 5' adaptor sequence: 5'
 GAATTCGGACGACGAG 3' 3' adaptor sequence: 5'
 CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.2 kb."

BASE COUNT 22 a 7 c 19 g 16 t

ORIGIN

Alignment Scores: 4.28e+03 Length: 64
 Pred. No.: 25.00 Matches: 4
 Score: 100.00% Conservativeness: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 100.00% Gaps: 0
 DB: 9

US-09-397-967A-17 (1-4) x AA595043 (1-64)

QY 1 LysAspTyrTyr 4

Db 41 AAGGATTATTAT 52

RESULT 11

BG656411/c 65 bp mRNA linear EST 05-JUL-2001
 LOCUS ib37b02.x1 HR85 islet Homo sapiens CDNA 3', mRNA sequence.
 DEFINITION BG656411
 ACCESSION BG656411.1 GI:13793820
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM

REFERENCE

AUTHORS Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
 Lemishka, I., Scearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 Hillier, L., Marra, M., Page, D., Wylie, T., Martin, J., Blistain, A.,
 Schmitt, A., Theising, B., Ritter, B., Ronko, I., Bennett, J., Cardenas
 , M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.,
 Jackson, Y. and Bowers, Y.
 Endocrine Pancreas Consortium
 Unpublished

TITLE

JOURNAL

COMMENT

Other ESTs: ib37b02.y1
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohpb.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Dr. Hiroshi Inoue
 (hinoue@im.wustl.edu)

Seq primer: -40UP from Gibco.

Location/Qualifiers

source

1. .65
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="Purified pancreatic islet"
 /lab_host="DH10B"
 /clone_lib="HR85 islet"
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:

NotI; Site 2: XhoI; cDNA made by oligo-dT priming.
 Size-selected on agarose gel. Average insert size ~1kb. 5'
 XhoI site was destroyed after directional cloning.
 Amplified once. Contact information: Hiroshi Inoue, MD,
 Metabolism Div. (Alan Permutt Lab), Washington University
 School of Medicine, Box 8127, 660 South Euclid Ave., St.
 Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
 314-362-1916, Fax: 314-747-2692."

BASE COUNT 25 a 5 c 9 g 26 t

ORIGIN

Alignment Scores: 4.36e+03 Length: 65
 Pred. No.: 25.00 Matches: 4
 Score: 100.00% Conservativeness: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 100.00% Gaps: 0
 DB: 10

US-09-397-967A-17 (1-4) x BG656411 (1-65)

QY 1 LysAspTyrTyr 4

Db 14 AAGGATTATTAT 3

RESULT 12

LOCUS

DEFINITION

AA388894 67 bp mRNA linear EST 23-APR-1997
 mp13h03.r1 Life Tech mouse embryo 8 5dpc 10664019 Mus musculus CDNA
 clone IMAGE:569141 5' similar to TT:G1235727 G1235727 SUPPRESSOR
 FOR YEAST MUTANT. ;, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 67)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Scheilenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 The WashU-HMI Mouse EST Project
 Unpublished

TITLE

JOURNAL

COMMENT

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:343789

Trace considered overall poor quality

Possible reversed clone: Similarity on wrong strand

Seq primer: -28m13 rev1 ET from Amersham

High quality sequence stop: 1.

Location/Qualifiers

source

1. .67
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:569141"
 /tissue_type="embryo"
 /dev_stage="8.5dpc embryos"
 /lab_host="DH10B"
 /clone_lib="Life Tech mouse embryo 8 5dpc 10664019"
 /note="Organ: whole embryo; Vector: pCMV-SPORT2; Site_1:
 SalI; Site_2: NotI; Cloned unidirectionally. Primer:

BASE COUNT 20 a 21 c 11 g 15 t

ORIGIN

Alignment Scores:

Pred. No.: 4.51e+03 Length: 67
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-397-967A-17 (1-4) x AA388894 (1-67)

Oy 1 LysAspTyrTyr 4

Db 26 AAAGACTACTAC 37

RESULT 13

BH792714

LOCUS 68 bp DNA linear GSS 02-APR-2002
DEFINITION SALK_064938.46.55.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_064938.46.55.x, genomic
survey sequence.

ACCESSION

VERSION BH792714.1 GI:19889800

KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM

REFERENCE 1 (bases 1 to 68)
AUTHORS Alonso,J.M., Karnes,M., Kim,C.J., Parker,H., Cheuk,R., Gadrinab
, C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.,
Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome

JOURNAL

COMMENT

Unpublished
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.

FEATURES

source

Location/Qualifiers

1..68
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"

/clone="SALK_064938.46.55.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT 19 a 11 c 9 g 29 t

ORIGIN

Alignment Scores:

Pred. No.: 4.59e+03 Length: 68
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 28 Gaps: 0

US-09-397-967A-17 (1-4) x BH792714 (1-68)

Oy 1 LysAspTyrTyr 4

Db 13 AAGGATTATTAT 24

RESULT 14

AL755459/c

LOCUS 69 bp DNA linear GSS 17-JUN-2002

DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-098A06-012117,
genomic survey sequence.

ACCESSION

VERSION AL755459.1 GI:21487957

KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM

REFERENCE 1 (bases 1 to 69)
AUTHORS Strizhov,N., Li,Y., Rosso,M., Viehoveer,P., Dekker,K., Saedler,H.
and Weisshaar,B.
TITLE A pipeline for automated high-throughput generation of FSTs
(flanking sequence tags) from Arabidopsis thaliana T-DNA
transformed lines
Unpublished

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
3 (bases 1 to 69)
Li,Y., Rosso,M., Strizhov,N. and Weisshaar,B.
Direct Submission
Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer
Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany

This sequence is recovered from the right border of the T-DNA. It
indicates an insertion within the locus defined by clone F12F1. The
sequences are generated at the MPI for Plant Breeding Research in
the context of the GABI-Kat project. GABI-Kat is part of the German
Plant Genomics program designated 'GABI'. Information on line
availability can be found at:
<http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES

source

Location/Qualifiers

1..69

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="GK-098A06-012117"

/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/note="PCR was performed on DNA from Arabidopsis thaliana
plants (T1) which were transformed with the T-DNA from
vector PAC161. The lines contain one or more T-DNA
insertions. The DNA fragment(s) resulting from the PCR
were directly sequenced to determine the genomic sequence
flanking the insertion. Sequences displaying significant
similarity to the A. thaliana nuclear genome sequence were
processed for submission. T-DNA derived sequences were
removed"

BASE COUNT 24 a 13 c 5 g 27 t

ORIGIN

Alignment Scores:

Pred. No.: 4.67e+03 Length: 69
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 29 Gaps: 0

US-09-397-967A-17 (1-4) x AL755459 (1-69)

QY 1 LysAspTyrTyr 4
 Db 16 AAAGATTATTAT 5

Db 54 AAAGATTACTAC 65

RESULT 15

AL764243

LOCUS

DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-123E01-012812, genomic survey sequence.

ACCESSION AL764243

VERSION AL764243.1

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1

AUTHORS

Strizhov, N., Li, Y., Rosso, M., Viehoveer, P., Dekker, K., Saedler, H. and Weisshaar, B.

TITLE

A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines

JOURNAL

REFERENCE 2

AUTHORS

Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weisshaar, B.

TITLE

A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics

JOURNAL

REFERENCE 3

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
 This sequence is recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At3g05630. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at:
<http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES

Source

1. .72 Location/Qualifiers

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="GK-123E01-012812"

/notes="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"

BASE COUNT 24 a 16 c 15 g 17 t

ORIGIN

Alignment Scores:

Pred. No.:	4.9e+03	Length:	72
Score:	25.00	Matches:	4
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	29	Gaps:	0

US-09-397-967A-17 (1-4) x AL764243 (1-72)

QY 1 LysAspTyrTyr 4

Search completed: February 4, 2004, 08:14:56
 Job time : 29.4633 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2004, 00:42:56 ; Search time 2.90481 Seconds
(without alignments)
3717.204 Million cell updates/sec

Title: US-09-397-967A-17
Perfect score: 25
Sequence: 1 KDYY 4

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlh
-Q=/cgn2 1/USPTO.spool/US09397967/runat_03022004.175631.28568/app_query.fasta_1.1486
-DB=N Geneseq 19Jun03 -OFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLAMP=200 -THR SCORE=100 -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HRAPISE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967 @CGN 1.1.501 @runat_03022004.175631.28568 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq 19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
c 1	25	100.0	12	23	ABI57762	Oligonucleotide pr
c 2	25	100.0	12	23	ABI78930	Oligonucleotide pr
c 3	25	100.0	13	23	ABC49776	Oligonucleotide SE
c 4	25	100.0	13	23	ABC49777	Oligonucleotide SE
c 5	25	100.0	22	16	AAQ85884	Primer for amplify
c 6	25	100.0	24	22	AAI69602	Human calthrin lig
c 7	25	100.0	25	21	AAA68498	Bacteriophage 3A O
c 8	25	100.0	29	21	AAQ4425	Polymorphic fragme
c 9	25	100.0	30	24	ABX69473	Novel Helicobacter
c 10	25	100.0	47	21	AAZ66789	Human map-related
c 11	25	100.0	51	22	AAI75504	Human silent SNP c
c 12	25	100.0	51	22	AAI75505	Human silent SNP c
c 13	25	100.0	63	21	AAC11379	Human secreted pro
c 14	25	100.0	64	13	AAQ33504	Downstream sequenc
c 15	25	100.0	71	22	AAF90751	Sequence #28 isola
c 16	25	100.0	76	21	AAA60873	Human secretory ph
c 17	25	100.0	94	25	ABZ78674	Tumour suppression
c 18	25	100.0	94	25	ABZ09221	Human oligonucleot
c 19	25	100.0	99	25	ABX08127	S. pneumoniae type
c 20	25	100.0	99	25	ABX48134	Bovine EST associa
c 21	25	100.0	107	16	AAI06096	Substance P bindin
c 22	25	100.0	108	20	AAZ33023	Humanised ATR-5 H
c 23	25	100.0	108	22	AAF87848	ATR-5 version a CD
c 24	25	100.0	110	20	AAZ33025	Humanised ATR-5 H
c 25	25	100.0	110	22	AAF87850	ATR-5 version a CD
c 26	25	100.0	119	22	AAK70919	Human immune/haema
c 27	25	100.0	122	21	AAC19685	Human secreted pro
c 28	25	100.0	123	21	AAC22090	Human secreted pro
c 29	25	100.0	138	22	AAZ07099	DNA encoding Group
c 30	25	100.0	150	22	AAI29196	Colon tumour relat
c 31	25	100.0	150	25	ABZ33382	Human colon tumour
c 32	25	100.0	153	22	ABA66324	Human foetal liver
c 33	25	100.0	153	22	AAK14742	Human brain expres
c 34	25	100.0	153	22	AAK40484	Human bone marrow
c 35	25	100.0	153	25	ACA00545	C. glutamicum deri
c 36	25	100.0	159	22	AAK74089	Human immune/haema
c 37	25	100.0	165	16	AAQ87512	Tyrosine kinase pe
c 38	25	100.0	177	21	AAC21103	Human secreted pro
c 39	25	100.0	177	24	ABN78876	Human ORF3823 CDNA
c 40	25	100.0	183	24	ABN85608	Human ovarian canc
c 41	25	100.0	184	22	AAK23938	Human brain expres
c 42	25	100.0	184	22	AAK50009	Human bone marrow
c 43	25	100.0	184	24	ABZ33501	Human genome-deriv
c 44	25	100.0	185	21	AAC22811	Human secreted pro
c 45	25	100.0	186	22	AAI61505	Soybean 240017 reg

ALIGNMENTS

RESULT 1
ABI57762/c
ID ABI57762 standard; DNA; 12 BP.
XX
XX ABI57762;
XX

DT 22-FEB-2002 (first entry)

XX Oligonucleotide primer SEQ ID NO 357735 for detecting SNP TSC0050754.

XX SNP: single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
XX peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
XX central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX OS
XX Homo sapiens.

XX WO200177384-A2.
EN

XX PS Claim 1; SEQ ID 49793; 29pp + Sequence Listing; German.

XX CC This invention describes novel oligonucleotide primers or peptide nucleic acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP) and cytosine methylation status in chemically pretreated genomic DNA. The CC CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a CC range of diseases including immune system, gastrointestinal, respiratory, CC central nervous system, cardiovascular and metabolic disorders. The CC oligomers are also used for detecting cell type differentiation. CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and CC ABH00010-ABH99989 represent the oligomers described in the invention. CC NOTE: The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic format from WIPO at CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 13 BP; 7 A; 0 C; 1 G; 5 T; 0 other;

Alignment Scores: Pred. No.: 137 Length: 13 Score: 25.00 Matches: 4 Percent Similarity: 100.00% Conservatively: 0 Best Local Similarity: 100.00% Mismatches: 0 Query Match: 100.00% Indels: 0 DB: 23 Gaps: 0

US-09-397-967A-17 (1-4) x ABC49776 (1-13)

QY 1 LysAspTyrTyr 4

Db 1 AAAGATTATTAT 12

RESULT 4

ABC49777/c

ID ABC49777 standard; DNA; 13 BP.

AC ABC49777;

XX 21-FEB-2002 (first entry)

DE Oligonucleotide SEQ ID NO 49794 for detecting SNP TSC0014050.

XX SNP: single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS; peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss; central nervous system; gastrointestinal; respiratory; immune; metabolic.

XX Homo sapiens.

XX WO200177384-A2.

XX 18-OCT-2001.

XX 06-APR-2001; 2001WO-IB00713.

XX 07-APR-2000; 2000DE-1019173.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2001-657177/75.

XX Set of oligonucleotides, useful for diagnosis and cell typing, is PT designed to detect single nucleotide polymorphisms and cytosine PT methylation status -

XX Claim 1; SEQ ID 49794; 29pp + Sequence Listing; German.

XX This invention describes novel oligonucleotide primers or peptide nucleic acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP) and cytosine methylation status in chemically pretreated genomic DNA. The CC CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a CC range of diseases including immune system, gastrointestinal, respiratory,

CC central nervous system, cardiovascular and metabolic disorders. The CC oligomers are also used for detecting cell type differentiation. CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and CC ABH00010-ABH99989 represent the oligomers described in the invention. CC NOTE: The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic format from WIPO at CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 13 BP; 5 A; 1 C; 0 G; 7 T; 0 other;

Alignment Scores: Pred. No.: 137 Length: 13 Score: 25.00 Matches: 4 Percent Similarity: 100.00% Conservatively: 0 Best Local Similarity: 100.00% Mismatches: 0 Query Match: 100.00% Indels: 0 DB: 23 Gaps: 0

US-09-397-967A-17 (1-4) x ABC49777 (1-13)

QY 1 LysAspTyrTyr 4

Db 13 AAAGATTATTAT 2

RESULT 5

AAQ85884

ID AAQ85884 standard; cDNA; 22 BP.

XX AAQ85884;

XX 25-MAR-2003 (updated)

DT 24-OCT-1995 (first entry)

XX Primer for amplifying gag gene of feline immunodeficiency virus.

XX Feline immunodeficiency virus; FIV; CD4; CD8; antigen; vaccine; diagnosis; primer; probe; acquired immune deficiency syndrome; AIDS; model system; screening; long terminal repeat; LTR; ss.

XX Synthetic.

XX WO9505460-A2.

XX 23-FEB-1995.

XX 25-JUL-1994; 94WO-US08364.

XX 12-AUG-1993; 93US-0105710.

XX 03-SEP-1991; 91US-0752424.

XX (UYNC-) UNIV NORTH CAROLINA STATE.

XX Tompkins MB, Tompkins WAF;

XX WPI; 1995-098760/13.

XX New isolate of feline immunodeficiency virus - and infected host PT cells, DNA, vectors and infected cats, useful in vaccines and as PT models for human AIDS

XX Example 12; Page 33; 50pp; English.

XX The NCSU strain (ATCC VR2333) of the feline immunodeficiency virus CC (FIV) is highly infectious in vivo and causes a rapid inversion of the CC CD4+CD8+ ratio in infected animals. Antigenic fragments of the CC virus can be used in vaccines. Fragments of the DNA can also be used CC diagnostically as probes and primers and for expressing viral CC antigens. Cats infected with the virus are useful as model systems CC for studying AIDS and for screening therapeutic drugs. Two primers CC (AAQ85883, AAQ85884) were used to amplify the gag gene of the NCSU CC strain of FIV (AAQ85885) for its subcloning and subsequent CC sequencing. NOTE: This sequence is the primer cited in the CC example although a second SEQ ID 5 appears in the sequence listing of

CC the specification. The second sequence is TCTGGCCTCA TTTCATATGA TG,
 CC this sequences complement.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX

SQ Sequence 22 BP; 10 A; 4 C; 5 G; 3 T; 0 other;

Alignment Scores:

Pred. No.: 229 Length: 22
 Score: 25.00 Matches: 4
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 16 Gaps: 0

US-09-397-967A-17 (1-4) x AAQ85884 (1-22)

QY 1 LysAspTyrTyr 4
 Db 11 AAAGATTACTAC 22

RESULT 6

ID AAI69602/c
 ID AAI69602 standard; DNA; 24 BP.

XX AC AAI69602;

XX 19-DEC-2001 (first entry)

DE Human calthrin light chain 17 coding sequence PCR primer #2.

XX Human; calthrin light chain 17; cancer; haemopathy; HIV infection;
 KW immunological disease; inflammation; gene therapy; PCR primer; ss.
 OS Homo sapiens.

XX WO200175045-A2.

PN 11-OCT-2001.

PF 26-MAR-2001; 2001WO-CN00456.

PR 28-MAR-2000; 2000CN-0115248.

XX (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.

PI Mao Y, Xie Y;

XX WPI; 2001-639358/73.

XX New human clathrin light chain 17 for diagnosing and treating malignant
 PT tumor, hemopathy, human immunodeficiency virus infection, immunological
 PT diseases and inflammation -
 XX Example 2; Page 16; 32pp; Chinese.

XX The present invention provides the protein and coding sequences of human
 CC calthrin light chain 17. The sequences can be used in the treatment of
 CC cancer, haemopathy, HIV infection, immunological diseases and
 CC inflammation. The present sequence is a PCR primer for the coding
 CC sequence of the invention.

SQ Sequence 24 BP; 7 A; 2 C; 2 G; 13 T; 0 other;

Alignment Scores:

Pred. No.: 249 Length: 24
 Score: 25.00 Matches: 4
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-397-967A-17 (1-4) x AAI69602 (1-24)

QY 1 LysAspTyrTyr 4
 Db 14 AAAGATTACTAC 3

RESULT 7

ID AAA68498 standard; DNA; 25 BP.

XX AC AAA68498;

XX 27-OCT-2000 (first entry)

XX Bacteriophage 3A ORF RBS sequence 3AORF244.

KW Bacteriophage; antimicrobial; genome; identification; antibacterial;
 KW bacterial growth inhibition; PCR primer; RBS; ribosome binding site;
 KW bacterial infection; ss.

OS Bacteriophage 3A.

PN WO200032825-A2.

XX 08-JUN-2000.

XX 03-DEC-1999; 99WO-IB02040.

XX 03-DEC-1998; 98US-0110992.

PR 03-JUN-1999; 99US-0328144.

PR 28-SEP-1999; 99US-0407804.

PR 30-SEP-1999; 99US-0157218.

PR 01-DEC-1999; 99US-0168777.

PR 02-DEC-1999; 99US-0454252.

XX (PHAG-) PHAGETECH INC.

XX Pelletier J, Gros P, Dubow M;

XX WPI; 2000-412361/35.

XX Identifying a bacteriophage coding region for treating bacterial

PT infections comprises identifying a nucleic acid encoding a product that
 PT inhibits bacteria when a bacteriophage infects a bacterium -
 XX Disclosure; Page 187; 456pp; English.

XX The present invention describes a method for identifying a bacteriophage
 CC coding region encoding a product active on an essential bacterial
 CC target. The method comprises identifying a nucleic acid sequence encoding
 CC a gene product that provides a bacteria-inhibiting function when an
 CC uncharacterised bacteriophage infects a pathogenic bacterium. The
 CC compound active on a target of a bacteriophage inhibitor protein in a
 CC bacteria is used to treat or prevent a bacterial infection in an animal.

CC AA68243 to AAA69442 and AAB16523 to AAB16954 represent bacteriophage
 CC nucleotide and protein sequences which are used in the exemplification of
 CC the present invention.

XX Sequence 25 BP; 10 A; 2 C; 2 G; 11 T; 0 other;

Alignment Scores:

Pred. No.: 259 Length: 25
 Score: 25.00 Matches: 4
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0

US-09-397-967A-17 (1-4) x AAA68498 (1-25)

QY 1 LysAspTyrTyr 4
 Db 11 AAAGATTACTAT 22

RESULT 8

AAA04425
 ID AAA04425 standard; DNA; 29 BP.
 AC AAA04425;
 XX
 XX 22-MAY-2000 (first entry)
 XX
 XX Polymorphic fragment of hypertension associated gene MRL.
 DE
 XX Polymorphism; hypertension; agammaglobulinemia; diabetes insipidus;
 KW Lesch-Nyhan syndrome; muscular dystrophy; Wiskott-Aldrich syndrome;
 KW Fabry's disease; familial hypercholesterolemia; hereditary spherocytosis;
 KW polycystic kidney disease; von Willebrand's disease; forensic; human;
 KW tuberculous sclerosis; hereditary hemorrhagica telangiectasia;
 KW familial colonic polyposis; hereditary hemorrhagica telangiectasia;
 KW Ehlers-Danlos syndrome; ss.
 XX
 OS Homo sapiens.
 XX
 XX EP955382-A2.
 PN
 XX 10-NOV-1999.
 PD
 XX 07-MAY-1999; 99EP-0250150.
 PF
 XX 07-MAY-1999; 98US-0084641.
 PR
 XX 03-MAY-1999; 99US-0304232.
 PR
 XX (AFFY-) AFFYMETRIX INC.
 PA (UYCA-) UNIV CASE WESTERN RESERVE.
 PA
 XX Fan JB, Chakravarti A, Haluska MK;
 PI
 XX WPI; 2000-107928/10.
 DR
 XX Novel nucleic acids containing polymorphisms used in the diagnosis of
 PT hypertension -
 PT
 XX
 XX Claim 1; Page 36; 53pp; English.
 PS
 XX The invention provides polymorphic fragments of genes associated with
 CC hypertension. The nucleic acids including the polymorphic sites can be
 CC used as probes or primers for expressing variant proteins. Detection of
 CC the polymorphisms is useful in designing prophylactic and therapeutic
 CC regimes customized to underlying abnormalities. The polymorphisms can be
 CC used for association studies for hypertension, and in hypertension
 CC diagnostic assays. Where the polymorphisms have strong correlation with
 CC hypertension, within a gene, they are likely to have a causative role in
 CC hypertension. This information can be used to find the precise role of a
 CC polymorphism in the disease, and this can be used to identify potential
 CC drugs which combat the disease. The polymorphisms can be tested for
 CC association with other diseases e.g. agammaglobulinemia, diabetes
 CC insipidus, Lesch-Nyhan syndrome, muscular dystrophy, Wiskott-Aldrich
 CC syndrome, Fabry's disease, familial hypercholesterolemia, polycystic
 CC kidney disease, hereditary spherocytosis, von Willebrand's disease,
 CC tuberculous sclerosis, hereditary hemorrhagica telangiectasia, familial
 CC colonic polyposis, Ehlers-Danlos syndrome, osteogenesis imperfecta, and
 CC acute intermittent porphyria. The polymorphic forms can also be used in
 CC forensics to identify individuals.
 XX
 SQ Sequence 29 BP; 10 A; 5 C; 5 G; 8 T; 1 other;
 Alignment Scores:
 Pred. No.: 299 Length: 29
 Score: 25.00 Matches: 4
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-397-967A-17 (1-4) x AAA04425 (1-29)
 QY 1 LysAspTyrTyr 4

Db 7 AAAGACTATAT 18
 RESULT 9
 ABX69473
 ID ABX69473 standard; DNA; 30 BP.
 XX
 XX AC ABX69473;
 XX
 XX 07-MAY-2003 (first entry)
 DT
 XX Novel Helicobacter pylori gene PCR primer #2444.
 DE
 XX Protein-protein interaction; ulcer; selected interacting domain;
 KW SID; PCR; primer; ss.
 KW
 XX Helicobacter pylori.
 OS
 XX WO200266501-A2.
 PN
 XX 29-AUG-2002.
 PD
 XX 28-DEC-2001; 2001WO-BP15428.
 PF
 XX 02-JAN-2001; 2001US-259302P.
 PR
 XX (HYBR-) HYBRIGENICS.
 PA (INSP) INST PASTEUR.
 PA
 XX Legrain P, Rain J, Colland F, De Reuse H, Labigne A;
 PI
 XX WPI; 2002-674910/72.
 DR
 XX New complexes of protein-protein interactions in Helicobacter pylori,
 PT useful for identifying modulating compounds for treating or preventing
 PT ulcers in mammals -
 PT
 XX Example 9; Page 564; 642pp; English.
 PS
 XX The invention describes a complex of protein-protein interactions in
 CC Helicobacter pylori selected from 421 complexes given in the
 CC specification. The complex of protein-protein interactions are useful
 CC for screening for agents which modulate the interaction of proteins.
 CC Modulating compounds which binds to a targeted bacterial protein may be
 CC used for treating or preventing ulcers in a human or animal. This
 CC sequence represents a primer used to isolate polynucleotides encoding
 CC Helicobacter pylori proteins for studies on protein-protein
 CC interactions.
 XX
 SQ Sequence 30 BP; 13 A; 4 C; 3 G; 7 T; 3 U; 0 other;
 Alignment Scores:
 Pred. No.: 309 Length: 30
 Score: 25.00 Matches: 4
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 US-09-397-967A-17 (1-4) x ABX69473 (1-30)
 QY 1 LysAspTyrTyr 4
 Db 17 AAAGATTACTAT 28
 RESULT 10
 AAZ66789/c
 ID AAZ66789 standard; DNA; 47 BP.
 XX
 XX AAZ66789;
 AC
 XX 10-SEP-2001 (first entry)
 DT
 XX

DE Human map-related biallelic marker SEQ ID NO:1136.
 XX Human genome; biallelic marker; high density disequilibrium map;
 XX genomic map; haplotype; phenotype; polymorphic base; genotyping;
 KW haplotyping; hybridisation; identification; characterisation;
 KW diagnosis; single nucleotide polymorphism; SNP; ds.
 XX
 OS Homo sapiens.
 XX
 RH Key Location/Qualifiers
 FT variation replace(24,G)
 FT /*tag= a
 FT /standard_name= "single nucleotide polymorphism"
 XX
 PN WO9954500-A2.
 XX
 PD 28-OCT-1999.
 XX
 PF 21-APR-1999; 99WO-IB00822.
 XX
 PR 21-APR-1998; 98US-0082614.
 XX
 PR 23-NOV-1998; 98US-0109732.
 XX
 XX (GEST) GENSET.
 XX
 XX Cohen D, Blumenfeld M, Chumakov I;
 PI WPI; 2000-013267/01.
 DR
 XX
 PT Novel biallelic markers used to construct a high density disequilibrium
 PT map of the human genome -
 XX
 PS Claim 1; Page 473; 2745pp; English.
 XX
 CC AA265654 to AA269578 represent human biallelic markers from the present
 CC invention, which contain a polymorphic base at position 24 of their
 CC nucleotide sequences. AA269579 to AA277440 represent amplification
 CC primers for the biallelic markers. The biallelic markers of the
 CC invention have a variety of uses: they can be used for high density
 CC mapping of the human genome, and in complex association studies and
 CC haplotyping studies which are useful in determining the genetic basis
 CC for disease states. Compositions and methods of the invention can also
 CC be useful for the identification of the targets for the development of
 CC pharmaceutical agents and diagnostic methods, as well as the
 CC characterisation of the differential efficacious responses to and side
 CC effects from pharmaceutical agents acting on a disease as well as other
 CC treatment.
 CC N.B. The SEQ ID NOS 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297
 CC and 3367, are not actually given a sequence in the Sequence Listing
 CC from the present invention.
 XX
 SQ Sequence 47 BP; 16 A; 7 C; 7 G; 17 T; 0 other;

Alignment Scores:
 Pred. No.: 478 Length: 47
 Score: 25.00 Matches: 4
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0

US-09-397-967A-17 (1-4) x AA266789 (1-47)
 OY 1 LysAspTyrTyr 4
 DB 47 AAGGATTACTAT 36

RESULT 11
 AAI75504/c
 ID AAI75504 standard; DNA; 51 BP.
 XX
 AC AAI75504;
 XX

DT 09-NOV-2001 (first entry)
 XX
 DE Human silent SNP containing nucleic acid SEQ:2445.
 XX
 KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
 KW protein therapy; vaccine; probe; diagnostic assay; detection;
 KW quantitation; restorative therapy; polymorphic; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200140521-A2.
 XX
 PD 07-JUN-2001.
 XX
 PF 30-NOV-2000; 2000WO-US32758.
 XX
 PR 30-NOV-1999; 99US-0168138.
 XX
 PR 29-NOV-2000; 2000US-0726173.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Shinkets RA, Leach M;
 XX
 XX WPI; 2001-356160/37.
 XX
 DR Polymorphic nucleic acid sequences, useful in genetic testing and
 PT therapy -
 PT
 XX
 PS Claim 1; Page 799; 2653pp; English.
 XX
 CC AA173060 to AA179867 represent isolated human polymorphic polynucleotide
 CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
 CC AA173114 to AA175329 represent peptides related to human polymorphic
 CC polynucleotide sequences. The sequences can be used in gene and protein
 CC therapy, and in vaccine production. (I) and the polypeptides encoded by
 CC them may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate expression of polymorphic polypeptides.
 CC For example, (I) may be used to treat disorders by rectifying mutations
 CC or deletions in a patient's genome that affect the activity of
 CC polypeptides by expressing inactive proteins or to supplement the
 CC patient's own production of polypeptide. Additionally, (I) and its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acids
 CC in samples, and therefore which patients may be in need of restorative
 CC therapy. The polypeptides encoded by (I) may be used as antigens in the
 CC production of antibodies specific for polymorphic polypeptides. The
 CC antibodies may also be used to down regulate expression and activity.
 CC The antibodies may also be used as diagnostic agents for detecting the
 CC presence of polymorphic polypeptides in samples.
 XX
 SQ Sequence 51 BP; 10 A; 11 C; 9 G; 21 T; 0 other;

Alignment Scores:
 Pred. No.: 517 Length: 51
 Score: 25.00 Matches: 4
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-397-967A-17 (1-4) x AA175504 (1-51)
 OY 1 LysAspTyrTyr 4
 DB 44 AAGGATTATTAC 33

RESULT 12
 AAI75505/c
 ID AAI75505 standard; DNA; 51 BP.
 XX
 AC AAI75505;
 XX
 DT 09-NOV-2001 (first entry)

XX DE Human silent SNP containing nucleic acid SEQ:2446.
 XX KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
 KW protein therapy; vaccine; probe; diagnostic assay; detection;
 KW quantitation; restorative therapy; polymorphic; ds.
 XX OS Homo sapiens.
 XX PN WO200140521-A2.
 XX PD 07-JUN-2001.
 XX PF 30-NOV-2000; 2000WO-US32758.
 XX PR 30-NOV-1999; 99US-0168138.
 XX PR 29-NOV-2000; 2000US-0726173.
 XX PA (CURA-) CURAGEN CORP.
 XX PI Shimkets RA, Leach M;
 XX PT WPI; 2001-356160/37.
 XX DR Polymorphic nucleic acid sequences, useful in genetic testing and
 PT therapy -
 PS Claim 1; Page 800; 2653pp; English.
 CC AA173060 to AA179867 represent isolated human polymorphic polynucleotide
 CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
 CC AA53114 to AA53329 represent peptides related to human polymorphic
 CC polynucleotide sequences. The sequences can be used in gene and protein
 CC therapy, and in vaccine production. (I) and the polypeptides encoded by
 CC them may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate expression of polymorphic polypeptides.
 CC For example, (I) may be used to treat disorders by rectifying mutations
 CC or deletions in a patient's genome that affect the activity of
 CC polypeptides by expressing inactive proteins or to supplement the
 CC patients own production of polypeptide. Additionally, (I) and its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acids
 CC in samples, and therefore which patients may be in need of restorative
 CC therapy. The polypeptides encoded by (I) may be used as antigens in the
 CC production of antibodies specific for polymorphic polypeptides. The
 CC antibodies may also be used to down regulate expression and activity.
 CC The antibodies may also be used as diagnostic agents for detecting the
 CC presence of polymorphic polypeptides in samples.
 XX SQ Sequence 51 BP; 10 A; 12 C; 9 G; 20 T; 0 other;
 Alignment Scores:
 Pred. No.: 517 Length: 51
 Score: 25.00 Matches: 4
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DE: 22 Gaps: 0
 US-09-397-967A-17 (1-4) x AA175505 (1-51)
 QY 1 LysAspTyrTyr 4
 Db 44 AAGGATTATTAC 33
 RESULT 13
 AAC11379/c
 ID AAC11379 standard; cDNA; 63 BP.
 XX AC AAC11379;
 XX AC AAC11379;
 XX DT 06-OCT-2000 (first entry)
 XX DE Bos taurus.

DE Human secreted protein 5' EST, SEQ ID NO: 15454.
 XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 XX OS Homo sapiens.
 XX PN EP1033401-A2.
 XX PD 06-SEP-2000.
 XX PF 21-FEB-2000; 2000EP-0200610.
 XX PR 26-FEB-1999; 99US-0122487.
 XX PA (GEST) GENSET.
 XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX DR WPI; 2000-500381/45.
 XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX PS Claim 1; SEQ ID 15454; 71pp + CD-ROM; English.
 CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.
 XX SQ Sequence 63 BP; 18 A; 9 C; 8 G; 28 T; 0 other;
 Alignment Scores:
 Pred. No.: 635 Length: 63
 Score: 25.00 Matches: 4
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DE: 21 Gaps: 0
 US-09-397-967A-17 (1-4) x AAC11379 (1-63)
 QY 1 LysAspTyrTyr 4
 Db 19 AAAGATTACTAC 8
 RESULT 14
 AAQ33604
 ID AAQ33604 standard; DNA; 64 BP.
 XX AC AAQ33604;
 XX AC AAQ33604;
 XX DT 25-MAR-2003 (updated)
 XX DT 02-FEB-1993 (first entry)
 XX DE Downstream sequence of microsatellite from clone AGLA296.
 XX KW PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;
 KW genetic mapping; traits; amplification; ss.
 XX XX Bos taurus.

XX WO213102-A1.
 XX PD 06-AUG-1992.
 XX PF 15-JAN-1992; 92WO-US00340.
 XX PR 15-JAN-1991; 91US-0642342.
 XX (GENM-) GENMARK.
 XX Georges M, Massey JM;
 XX WPI; 1992-284684/34.
 XX Polymorphic bovine DNA markers - used in genetic identification,
 PT gene mapping, and selective breeding
 XX Table 7; Page 171; 517pp; English.
 XX The sequence is that downstream of a bovine microsatellite sequence
 CC obtd. by screening a library of bovine MboI DNA fragments of between
 CC 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.
 CC One out of 50 clones cross-hybridised. Assuming independent
 CC distribution of microsatellites and MboI sites, the frequency of
 CC (76)n >9 microsatellites in the bovine genome is estimated at >100,
 CC 000. The sequence information for ca. 230 such bovine microsatellites
 CC is summarised in the specification and indexed herein (see below).
 CC The sequences upstream and downstream of the microsatellite sequence
 CC were used to generate the required PCR primers for in vitro
 CC amplification of the corresp. microsatellite (using the program
 CC OPTIPRIM). The microsatellites may be used to identify individuals,
 CC for parentage testing, and in the genetic mapping of economic trait
 CC loci, or genes involved in the determination of economically important
 CC traits esp. in cattle, to allow selective breeding.
 CC See also AAQ33501-34437.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX SQ Sequence 64 BP; 25 A; 10 C; 12 G; 17 T; 0 other;

Alignment Scores:
 Pred. No.: 645 Length: 64
 Score: 25.00 Matches: 4
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0

US-09-397-967A-17 (1-4) x AAQ33604 (1-64)

QY 1 LysAspTyrTyr 4
 Db 20 AAGGATTATTAC 31

RESULT 15
 AAF90751/c
 ID AAF90751 standard; RNA; 71 BP.
 XX AAF90751;
 XX DT 03-MAY-2001 (first entry)
 XX DE
 XX DE Sequence #28 isolated from round 8 of surface plasmon resonance SELEX.
 XX KW Human; transforming growth factor beta2; TGFbeta2; SELEX;
 XX KW systemic evolution of ligands by exponential enrichment; ss.
 XX OS Homo sapiens.
 XX PN WO200109156-A1.
 XX PD 08-FEB-2001.

PF 26-JUL-2000; 2000WO-US20397.
 XX PR 29-JUL-1999; 99US-0363939.
 XX PA (NEXS-) NEXSTAR PHARM INC.
 XX PI Pagratis N, Lochrie M, Gold L;
 XX DR WPI; 2001-218217/22.
 XX New RNA ligand to human transforming growth factor beta2, useful as
 PT pharmaceuticals, diagnostics and as immunochemical reagents -
 XX Claim 1; Page 61; 178pp; English.
 XX The present invention relates to non-naturally occurring, high-affinity
 CC RNA ligands to human transforming growth factor beta2 (TGFbeta2). The
 CC oligonucleotide ligands were identified by the SELEX method (SELEX stands
 CC for Systemic Evolution of Ligands by Exponential Enrichment). The
 CC oligonucleotide ligands are useful in any process in which binding to
 CC TGFbeta2 is required. The ligands may be useful as pharmaceuticals,
 CC diagnostics, imaging agents and immunochemical reagents. The present
 CC sequence is an oligonucleotide used in the present invention.
 XX SQ Sequence 71 BP; 12 A; 22 C; 16 G; 21 U; 0 other;

Alignment Scores:
 Pred. No.: 713 Length: 71
 Score: 25.00 Matches: 4
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-397-967A-17 (1-4) x AAF90751 (1-71)

QY 1 LysAspTyrTyr 4
 Db 26 AAGGATTATTAC 15

Search completed: February 4, 2004, 03:08:18
 Job time : 3.90481 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2004, 13:04:22 ; Search time 873 Seconds

(without alignments)
4637.242 Million cell updates/sec

Title: US-09-397-967A-16

Perfect score: 1099

Sequence: 1 MAPSEETLPQPSCLSS.....RPAPATLSQLDPLWRGPG 1099

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2449703 seqs, 184161367 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4891994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Model line parameters:
-O=/cgn2_1/USPto_spool/US09397967/runat_03022004_175718_28991/app_query.fasta.1.1287
-DB=Published Applications NA -QFMT=fastap -SUFFIX=olip2n.rnpb -MINMATCH=0.1
-LOOPCU=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-ALIGN=15 -MODE=LOCAL -OUTFMT=PTO -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09397967 @CGN 1 1 331 @runat_03022004_175718_28991
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	62	5.6	3620	12	US-10-305-720-1055	Sequence 1055, Ap
2	62	5.6	5118	13	US-09-814-353-19356	Sequence 13556, A
3	58	5.3	203	9	US-09-728-446-1413	Sequence 1413, Ap
4	46	4.2	439	11	US-09-918-995-11571	Sequence 11571, A
5	28	2.5	438	11	US-09-918-995-10879	Sequence 10879, A
6	21	1.9	65	13	US-09-908-975-30425	Sequence 30425, A
7	20	1.8	860	14	US-10-027-632-8229	Sequence 8229, Ap
8	20	1.8	860	14	US-10-027-632-8229	Sequence 8229, Ap
9	18	1.6	364	10	US-09-771-161A-17	Sequence 17, Appl
10	16	1.5	3731	12	US-10-191-803-200	Sequence 200, App
11	16	1.5	4078	12	US-10-305-720-297	Sequence 297, App
12	16	1.5	7126	12	US-10-062-674-2214	Sequence 2214, Ap
13	12	1.1	151	10	US-09-982-610-11	Sequence 11, Appl
14	12	1.1	215	13	US-10-029-386-26806	Sequence 26806, A
15	12	1.1	294	9	US-09-864-761-21086	Sequence 21086, A
16	12	1.1	367	9	US-09-864-761-4335	Sequence 4335, Ap
17	12	1.1	372	10	US-09-878-574-1581	Sequence 1581, Ap
18	12	1.1	477	11	US-09-918-995-2066	Sequence 2066, Ap
19	12	1.1	559	15	US-10-066-543-1870	Sequence 1870, Ap
20	12	1.1	562	13	US-10-029-386-13106	Sequence 13106, A
21	12	1.1	637	13	US-10-027-632-247078	Sequence 247078, A
22	12	1.1	637	14	US-10-027-632-247078	Sequence 247078, A
23	12	1.1	917	11	US-09-991-936-1577	Sequence 1577, Ap
24	12	1.1	927	13	US-10-430-797-3	Sequence 3, Appl
25	12	1.1	1509	13	US-10-205-194-10	Sequence 10, Appl
26	12	1.1	1583	11	US-09-971-708-6	Sequence 6, Appl
27	12	1.1	1713	13	US-10-187-900-1	Sequence 1, Appl
28	12	1.1	1987	13	US-10-100-217-1	Sequence 1, Appl
29	12	1.1	1989	15	US-10-084-817-341	Sequence 341, App
30	12	1.1	2000	9	US-09-977-269-1	Sequence 1, Appl
31	12	1.1	2000	10	US-09-977-260-1	Sequence 1, Appl
32	12	1.1	2000	11	US-09-977-261-1	Sequence 1, Appl
33	12	1.1	2065	13	US-10-103-380A-1	Sequence 1, Appl
34	12	1.1	2829	13	US-10-430-797-1	Sequence 1, Appl
35	12	1.1	2982	13	US-10-187-958-2	Sequence 2, Appl
36	12	1.1	3137	10	US-09-954-456-1128	Sequence 1128, Ap
37	12	1.1	3136	13	US-10-412-277-1	Sequence 1, Appl
38	12	1.1	3170	13	US-09-971-392-66	Sequence 66, Appl
39	12	1.1	3220	13	US-10-430-797-5	Sequence 5, Appl
40	12	1.1	3348	10	US-09-982-610-34	Sequence 34, Appl
41	12	1.1	3370	10	US-09-967-768A-144	Sequence 144, App
42	12	1.1	3370	13	US-10-354-358-101	Sequence 101, App
43	12	1.1	3370	13	US-10-210-120-19	Sequence 19, Appl
44	12	1.1	3393	9	US-09-922-138-3	Sequence 3, Appl
45	12	1.1	3583	13	US-10-237-496-103	Sequence 103, App

ALIGNMENTS

RESULT 1
US-10-305-720-1055
; Sequence 1055, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; PRIORITY FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1055
; LENGTH: 3620
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank ID No. US20040010136A1 g1039418
US-10-305-720-1055

Alignment Scores:

Pred. No.: 9,75e-52 Length: 3620
 Score: 62.00 Matches: 96
 Percent Similarity: 97.96% Conservatives: 0
 Best Local Similarity: 97.96% Mismatches: 1
 Query Match: 5.64% Indels: 2
 DB: 12 Gaps: 0

US-09-397-967A-16 (1-1099) x US-10-305-720-1055 (1-3620)

QY 938 ArgArgCysValHisArgAspLeuAlaAlaArgAenlleLeuValGluSerGluAlaHis 957
 DB 2922 CGCCGCTGCTGCACCGGACCTGCGCCGCGAATCATCTCTGAGAGCGAGGACAC 2981
 QY 958 VallysilleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrVa 977
 DB 2982 GTCAGATCGCTGACTTCGCGCTAGCTAAGCTGCGCCGCTGAC-AAAGACTACTACGT 3040
 QY 977 lVallArgGluProGlyGlnSerProillePheTyrAlaProGluSerLeuSerAspAs 997
 DB 3041 GGTCCGCGAGCCAGCCAGCCCATTTCTGGTATGCCCGAATCCCTCTCGGACAA 3100
 QY 997 nilePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheTh 1017
 DB 3101 CATCTTCTCTCGCCAGTCAGAGCTCTGGAGCTTCGGGGTCTGCTGTACGAGCTCTTCAC 3160
 QY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034
 DB 3161 CTACTGCGACAAAGCTGAGCCCTCGCGGAGTTCTTGGCGATGATGGGA 3212

RESULT 2

US-09-814-353-19356

Sequence 19356, Application US/09814353
 Publication No. US20030165831A1
 GENERAL INFORMATION:
 APPLICANT: Thompson, Pamela
 APPLICANT: Lillie, James
 TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
 TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
 FILE REFERENCE: MRI-006B
 CURRENT APPLICATION NUMBER: US/09/814,353
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: US 60/191,031
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: US 60/207,124
 PRIOR FILING DATE: 2000-05-25
 PRIOR APPLICATION NUMBER: US 60/211,940
 PRIOR FILING DATE: 2000-06-15
 PRIOR APPLICATION NUMBER: US 60/216,820
 PRIOR FILING DATE: 2000-07-07
 PRIOR APPLICATION NUMBER: US 60/220,661
 PRIOR FILING DATE: 2000-07-25
 PRIOR APPLICATION NUMBER: US 60/257,672
 PRIOR FILING DATE: 2000-12-21
 NUMBER OF SEQ ID NOS: 22037
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 19356
 LENGTH: 5118
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 1, 2, 3, 4, 5114, 5115, 5116, 5117, 5118
 OTHER INFORMATION: n = A,T,C or G

US-09-814-353-19356

Alignment Scores:
 Pred. No.: 1.33e-51 Length: 5118
 Score: 62.00 Matches: 96
 Percent Similarity: 97.96% Conservatives: 0
 Best Local Similarity: 97.96% Mismatches: 1
 Query Match: 5.64% Indels: 2

DB: 13 Gaps: 0

US-09-397-967A-16 (1-1099) x US-09-814-353-19356 (1-5118)

QY 938 ArgArgCysValHisArgAspLeuAlaAlaArgAenlleLeuValGluSerGluAlaHis 957
 DB 2923 CGCCGCTGCTGCACCGGACCTGCGCCGCGAATCATCTCTGAGAGCGAGGACAC 2982
 QY 958 VallysilleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrVa 977
 DB 2983 GTCAGATCGCTGACTTCGCGCTAGCTAAGCTGCGCCGCTGAC-AAAGACTACTACGT 3041
 QY 977 lVallArgGluProGlyGlnSerProillePheTyrAlaProGluSerLeuSerAspAs 997
 DB 3042 GGTCCGCGAGCCAGCCAGCCCATTTCTGGTATGCCCGAATCCCTCTCGGACAA 3101
 QY 997 nilePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheTh 1017
 DB 3102 CATCTTCTCTCGCCAGTCAGAGCTCTGGAGCTTCGGGGTCTGCTGTACGAGCTCTTCAC 3161
 QY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034
 DB 3162 CTACTGCGACAAAGCTGAGCCCTCGCGGAGTTCTTGGCGATGATGGGA 3213

RESULT 3

US-09-728-446-1413

Sequence 1413, Application US/09728446
 Patent No. US20020081668A1
 GENERAL INFORMATION:
 APPLICANT: Friedrich, Glenn
 APPLICANT: Zambrowicz, Brian
 APPLICANT: Sands, Arthur T.
 TITLE OF INVENTION: No. US20020081668A1e1 Murine Polynucleotide Sequences
 TITLE OF INVENTION: and Mutant Cells and Mutant Animals Defined Thereby
 FILE REFERENCE: LEX-0101-USA
 CURRENT APPLICATION NUMBER: US/09/728,446
 CURRENT FILING DATE: 2000-11-30
 PRIOR APPLICATION NUMBER: US 60/168,270
 PRIOR FILING DATE: 1999-12-01
 NUMBER OF SEQ ID NOS: 1461
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1413
 LENGTH: 203
 TYPE: DNA
 ORGANISM: Mus musculus

US-09-728-446-1413

Alignment Scores:
 Pred. No.: 8.04e-49 Length: 203
 Score: 58.00 Matches: 58
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.28% Indels: 0
 DB: 9 Gaps: 0

US-09-397-967A-16 (1-1099) x US-09-728-446-1413 (1-203)

QY 1 MetAlaProProSerGluGluThrProLeuLeuProGlnArgSerCysSerLeuSerSer 20
 DB 18 ATGGCACCTCCAAAGTGGAGAGACCTCTGATCCCTCAGCGCTTTGACGCTCTCATCC 77
 QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40
 DB 78 TCAGAGCGAGGAGCCCTGCGATGTGCTCTCTCCCTCCCGGGGACCTGGCGCTCCCGACGGA 137
 QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAa 58
 DB 138 TTGTCAATCTCTTTTGGGGACTACTTGGCTGAGGATTTATGTGCGAGCTGCC 191

RESULT 4

US-09-918-995-11571

Sequence 11571, Application US/09918995
 Publication No. US20030073623A1

; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11571
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-11571

Alignment Scores:
Pred. No.: 1.92e-36 Length: 439
Score: 46.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.19% Indels: 0
DB: 11 Gaps: 0

US-09-397-967a-16 (1-1099) x US-09-918-995-11571 (1-439)

Qy 989 TyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTIpSerPhe 1008
Db 39 TATGCCCCCGAATCCCTCGGACAACTCTCTCTGCCAGTCAGAGCTTGAGCTTC 98
Qy 1009 GlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAlaGlu 1028
Db 99 GGGGTGCTGCTGTACGAGCTCTTCCACTACTACGACAAAAGCTGCAGCCCTCGGCCGAG 158
Qy 1029 PheLeuArgMetMetGly 1034
Db 159 TTCCTGCGGATGATGGGA 176

RESULT 5
US-09-918-995-10879
; Sequence 10879, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10879
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(438)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-10879

Alignment Scores:
Pred. No.: 2.54e-18 Length: 438
Score: 28.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.55% Indels: 0
DB: 11 Gaps: 0
US-09-397-967a-16 (1-1099) x US-09-918-995-10879 (1-438)

Qy 1007 SerPheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSer 1026
Db 96 AGCTTCGGGGTCTGCTGTACGAGCTCTTCCACTACTACGACAAAAGCTGCAGCCCTCG 155
Qy 1027 AlaGluPheLeuArgMetMetGly 1034
Db 156 GCCGAGTTCCTGCGGATGATGGGA 179

RESULT 6

US-09-908-975-30425
; Sequence 30425, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30425
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-30425

Alignment Scores:
Pred. No.: 5.2e-12 Length: 65
Score: 21.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.91% Indels: 0
DB: 13 Gaps: 0
US-09-397-967a-16 (1-1099) x US-09-908-975-30425 (1-65)

Qy 739 LeuProAlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAsp 758
Db 3 CTGCCCGCTCTCAAAATGACAGAACTGCGGGGACTTATCACAGTGCATGGCGTATGAT 62
Qy 759 Pro 759
Db 63 CCT 65

RESULT 7

US-10-027-632-8229/c
; Sequence 8229, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/219,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20

```

; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8229
; LENGTH: 860
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(860)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-8229

```

```

Alignment Scores:
Pred. No.: 5,23e-10 Length: 860
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.82% Indels: 0
DB: 13 Gaps: 0

```

US-09-397-967A-16 (1-1099) x US-10-027-632-8229 (1-860)

```

QY 889 ProGlyArgGlnSerLeuArgLeuValMetGluTyLeuProSerGlyCysLeuArgAsp 908
Db 366 CCAGGCCCGCAGAGCCTGGCGTGGTATGAGTACTGCCCGCAGCGCTGCTTGGCGGAC 307

```

```

RESULT 8
US-10-027-632-8229/c
; Sequence 8229, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8229
; LENGTH: 860
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(860)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-8229
Alignment Scores:

```

```

Pred. No.: 5,23e-10 Length: 860
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.82% Indels: 0
DB: 14 Gaps: 0
US-09-397-967A-16 (1-1099) x US-10-027-632-8229 (1-860)
QY 889 ProGlyArgGlnSerLeuArgLeuValMetGluTyLeuProSerGlyCysLeuArgAsp 908
Db 366 CCAGGCCCGCAGAGCCTGGCGTGGTATGAGTACTGCCCGCAGCGCTGCTTGGCGGAC 307

```

RESULT 9

```

US-09-771-161A-17
; Sequence 17, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 17
; LENGTH: 364
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-771-161A-17

```

```

Alignment Scores:
Pred. No.: 2,51e-08 Length: 364
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.64% Indels: 0
DB: 10 Gaps: 0

```

US-09-397-967A-16 (1-1099) x US-09-771-161A-17 (1-364)

```

QY 1 MetAlaProSerGluThrProLeuIleProGlnArgSerCysSerLeu 18
Db 96 ATGGCACCTCCAGTGAAGAGAGCGCCCTGATCCCTCAGCGTTTCATGCAGCCTC 149

```

RESULT 10

```

US-10-191-803-200
; Sequence 200, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10/191,803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611

```

FILE REFERENCE: 2002-05-06
CURRENT APPLICATION NUMBER: US/10/062,674
CURRENT FILING DATE: 2002-01-30
PRIOR APPLICATION NUMBER: US 09/625,102
PRIOR FILING DATE: 2000-07-24
NUMBER OF SEQ ID NOS: 2217
SOFTWARE: PERL Program
SEQ ID NO 2214
LENGTH: 7126
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040005559A1 903485.2
FEATURE:
NAME/KEY: unsure
LOCATION: (1) ... (7126)
OTHER INFORMATION: a, t, c, g, or other
US-10-062-674-2214

Alignment Scores:
Pred. No.: 2,05e-05 Length: 3731
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 12 Gaps: 0

US-09-397-967A-16 (1-1099) x US-10-191-803-200 (1-3731)

QY 1003 SerAspValTTPSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
DB 3259 TCAGATGTTGGAGCTTTGGTGTCTGTATGAACATTTTCACGTAT 3306

RESULT 11

US-10-305-720-297
Sequence 297, Application US/10305720
Publication No. US20040010136A1
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
FILE REFERENCE: PA-0002-1 CON
CURRENT APPLICATION NUMBER: US/10/305,720
CURRENT FILING DATE: 2002-11-26
PRIOR APPLICATION NUMBER: 09/016,434
PRIOR FILING DATE: 1998-01-30
NUMBER OF SEQ ID NOS: 1490
SOFTWARE: PERL Program
SEQ ID NO 297
LENGTH: 4078
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040010136A1 179527
FEATURE:
NAME/KEY: unsure
LOCATION: (1) ... (4078)
OTHER INFORMATION: a, t, c, g, or other
US-10-305-720-297

Alignment Scores:
Pred. No.: 2,22e-05 Length: 4078
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 12 Gaps: 0

US-09-397-967A-16 (1-1099) x US-10-305-720-297 (1-4078)

QY 1003 SerAspValTTPSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
DB 3441 TCAGATGTTGGAGCTTTGGTGTCTGTATGAACATTTTCACATAC 3488

RESULT 12

US-10-062-674-2214
Sequence 2214, Application US/10062674
Publication No. US20040005559A1
GENERAL INFORMATION:
APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS

FILE REFERENCE: PA-0026-1 CIP
CURRENT APPLICATION NUMBER: US/10/062,674
CURRENT FILING DATE: 2002-01-30
PRIOR APPLICATION NUMBER: US 09/625,102
PRIOR FILING DATE: 2000-07-24
NUMBER OF SEQ ID NOS: 2217
SOFTWARE: PERL Program
SEQ ID NO 2214
LENGTH: 7126
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040005559A1 903485.2
FEATURE:
NAME/KEY: unsure
LOCATION: (1) ... (7126)
OTHER INFORMATION: a, t, c, g, or other
US-10-062-674-2214

Alignment Scores:
Pred. No.: 3,64e-05 Length: 7126
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 12 Gaps: 0

US-09-397-967A-16 (1-1099) x US-10-062-674-2214 (1-7126)

QY 1003 SerAspValTTPSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
DB 3606 TCAGATGTTGGAGCTTTGGTGTCTGTATGAACATTTTCACATAC 3653

RESULT 13

US-09-982-610-11
Sequence 11, Application US/09982610
Patent No. US20020146420A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
Bennett, Brian D.
Goeddel, David
Lee, James M.
Matthews, William
Tsai, Siao Ping
Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982,610
FILING DATE: 17-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,648
FILING DATE: 1996-MAY-23
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0821P3PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
LENGTH: 151 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-982-610-11

Alignment Scores:
Pred. No.: 0.0126 Length: 151
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967A-16 (1-1099) x US-09-982-610-11 (1-151)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 1 GTGCACAGGATCTCGGGCTCGAATCTCGTC 36

RESULT 14

US-10-029-386-26806
; Sequence 26806, Application US/10029386
; Publication No. US20030194704A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: AEOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 26806
LENGTH: 215
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR3.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.69
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.58
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.35
OTHER INFORMATION: SWISSPROT HIT: Q07494, EVALUATE 2.00e-35
OTHER INFORMATION: NT HIT: AF037331.1, EVALUATE 1.00e-117
OTHER INFORMATION: EST_HUMAN HIT: BG481991.1, EVALUATE 2.00e-39
US-10-029-386-26806

Alignment Scores:
Pred. No.: 0.0173 Length: 215
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 13 Gaps: 0

US-09-397-967A-16 (1-1099) x US-10-029-386-26806 (1-215)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 88 GTGCATCGGACCTCGGCTCGTAGAACATCTCGTC 123

RESULT 15

US-09-864-761-21086
; Sequence 21086, Application US/09864761

Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Hanzel, David K.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: AEOMICA-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,587
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 21086
LENGTH: 294
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC010899.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.6
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 3.8
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.8
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.4
OTHER INFORMATION: SWISSPROT HIT: Q91694, EVALUATE 2.00e-35
OTHER INFORMATION: NT HIT: G14758279, EVALUATE 1.00e-115
OTHER INFORMATION: EST_HUMAN HIT: AW965297.1, EVALUATE 2.00e-26
US-09-864-761-21086

Alignment Scores:
Pred. No.: 0.0228 Length: 294
Score: 12.00 Matches: 12

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2004, 00:50:11 ; Search time 11745.3 Seconds
(without alignments)
3827.903 Million cell updates/sec

Title: US-09-397-967A-16
Perfect score: 5860
Sequence: 1 MAPSEETPLPQRSCSLSS.....RPAFATLSPQLDPLWRGRPG 1099

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV.xlh
-Q=/cgn2_1/USPTO_spool/US09397967/runat_03022004_175631_28576/app_query.fasta_1.1486
-DB=GenEmbl -QMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OJTPMT=pt0 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967@cgn_1_17598@runat_03022004_175631_28576 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_in.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_hgt_hum.*
31: em_hgt_inv.*
32: em_hgt_other.*
33: em_hgt_mus.*
34: em_hgt_pln.*
35: em_hgt_rod.*
36: em_hgt_mam.*
37: em_hgt_vrt.*
38: em_sy.*
39: em_hgt_hum.*
40: em_hgt_mus.*
41: em_hgt_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	5854	99.9	3723	10	MUSPTYKIN	L32955 Mouse prote
2	5687.5	97.1	3528	10	MUSJAK3A	L40172 Mus musculu
3	5380.5	91.8	3778	10	RATJAK3	D28508 Rattus norv
4	5223	89.1	4016	10	MUSJAK3H	L33788 Mus musculu
5	4684	79.9	4064	6	AX203081	AX203081 Sequence
6	4684	79.9	4064	9	HSU09607	U09607 Human JAK f
7	4597	78.4	3620	6	AR270492	AR270492 Sequence
8	4597	78.4	3620	9	HSU31601	U31601 Human tyros
9	4091.5	69.8	218208	2	AC073700	AC073700 Mus muscu
10	4091.5	69.8	223734	2	AC073750	AC073750 Mus muscu
11	3935	67.2	10326	10	MMU71201	U71201 Mus rattus
12	3692	63.0	216847	2	AC103459	AC103459 Rattus no
13	3336	56.9	3644	5	AF034576	AF034576 Gallus ga
14	3228.5	55.1	2914	6	AX203084	AX203084 Sequence
15	2713	45.3	13562	9	HSU70065	U70065 Human JAK3
16	2637.5	45.0	4812	4	AB006011	AB006011 Sus scrof
17	2636.5	45.0	3731	10	RN013396	U13396 Rattus norv
18	2629	44.9	3399	6	AX695805	AX695805 Sequence
19	2629	44.9	3435	6	AR129823	AR129823 Sequence
20	2629	44.9	4161	9	AF005216	AF005216 Homo sapi
21	2629	44.9	4216	9	BC039695	BC039695 Homo sapi
22	2629	44.9	5097	6	AX695804	AX695804 Sequence
23	2628	44.8	3500	9	AF001362	AF001362 Homo sapi
24	2621	44.7	5117	6	AX354568	AX354568 Sequence
25	2621	44.7	5117	9	AF058925	AF058925 Homo sapi
26	2619.5	44.7	3629	6	AR136148	AR136148 Sequence
27	2619.5	44.7	3629	6	AR143945	AR143945 Sequence
28	2619.5	44.7	3629	6	192575	192575 Sequence 8
29	2613.5	44.6	3390	6	AX695802	AX695802 Sequence
30	2613.5	44.6	3629	6	AX695801	AX695801 Sequence
31	2613.5	44.6	3629	10	MUSPTK	L16956 Mouse prote
32	2606.5	44.5	4387	5	DRE5690	AJ005690 Danio rer
33	2594	44.3	4078	6	AR269734	AR269734 Sequence
34	2523.5	43.1	3739	5	AF148993	AF148993 Cyprinus
35	2511	42.8	4042	4	AB036337	AB036337 Sus scrof
36	2460.5	42.0	2881	9	BC028068	BC028068 Homo sapi
37	2382.5	40.7	41006	9	AC007201	AC007201 Homo sapi
38	2360	40.3	3495	6	AR048200	AR048200 Sequence
39	2360	40.3	3495	6	AR068116	AR068116 Sequence
40	2360	40.3	3495	6	AR071176	AR071176 Sequence
41	2360	40.3	3495	6	AR177639	AR177639 Sequence
42	2360	40.3	3495	6	162349	162349 Sequence 2
43	2360	40.3	3495	6	187981	187981 Sequence 2
44	2350	40.1	4599	4	AB036336	AB036336 Sus scrof
45	2339	39.9	21323	9	AF513860	AF513860 Homo sapi

ALIGNMENTS

RESULT 1

```

MUSPTYKIN      3723 bp      mRNA      linear      ROD 21-SEP-1994
LOCUS
DEFINITION      Mouse protein tyrosine kinase (jak3) mRNA, complete cds.
ACCESSION      L32955
VERSION      L32955.1
KEYWORDS      Gi:529238
               interleukin 2; interleukin 4; Janus kinase; protein tyrosine
SOURCE
ORGANISM      Mus musculus (house mouse)
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS      Witthuhn, B.A., Silvennoinen, O., Miura, O., Lai, K.S., Cwik, C.,
               Liu, E.T. and Ihle, J.N.
TITLE      Involvement of the Jak-3 Janus kinase in signalling by interleukins
               2 and 4 in lymphoid and myeloid cells
JOURNAL      Nature 370 (6485), 153-157 (1994)
MEDLINE      94294024
PUBMED      8022486
COMMENT      Original source text: Mus musculus (strain BALB/c, sub_species
               domesticus) cDNA to mRNA.
               This entry has been reported under the accession number L329555 in
               the Nature article.
FEATURES
source
1..3723
   /organism="Mus musculus"
   /mol_type="mRNA"
   /strain="BALB/c"
   /sub_species="domesticus"
   /db_xref="taxon:10090"
   /cell_line="Abelson transformed pre-B-cell line 22D6"
   /cell_type="lymphoid and myeloid"
1..3723
   /gene="jak3"
1..12
   /gene="jak3"
13..3312
   /gene="jak3"
   /standard_name="Janus kinase, mujak3"
   /codon_start=1
   /evidence=experimental
   /product="protein tyrosine kinase"
   /protein_id="AAA21565.1"
   /db_xref="GI:529239"
   /translation="MAPPSETPLIPORCSLSSBAGALHVLPPRGPPORLSFS
               FGYLADLICVRAKACAILPVHSLPALATEDPFCWPPSHIFCIEDVDVQLVYRL
               RFYPPWFGLETCHRFGLKDLTSAIDLHLVSHLPAQRSDLVSGRLPVGLSMKEQG
               EFUSLAVLDLAQWARQAQRPGLKTVSYKACLPSLRDVIQGNFVTRRRIRRTVW
               LALLPCGRLPFRPYALMAKYILDLERLHPAATETTFVGLPGAQEEFGLRVAGDNGI
               PWSNDELFTQDFPIVDVSIQAPRVPGAGEHLVTVTRMDGHILAEFPGLPEA
               LSFVALVDGYFRLICDSRHVFCKEVAPPRLLEEDVCHGPITLDEAHLKAAGSLP
               GTYLRRSPQDYSPLTACTVQTPGLPDYKGLIRQDPGAFSLVGLSQPHRSRLRELL
               AACWSGLRVDAALYLTSCAPRPKEKSNLIVRRGDNPPAPGAPGSCCALQUSF
               HTTPTDSLEHNLHGHSFTKIPGRSERREVDTGTHDSEVLKVMDSRHRNCMESFLE
               AASLMQSVYPLHLVGLGCMAGDSIMVQEFVYLGAIDMYLRKRGHLVSAWSKLQVTK
               OLAVALNLYDKLGHNGVSARKVLAREGGDGNPPFIKLSDPGVSTPVLSEMLTDR
               IPWVAPCEAOATCLLEADKFGGATWEVFORGAHITSLRPAKKLKFYEDQGLP
               ALKWTLAGLITQCMAYDPGRSPFRALINDNLGLTSDYELLSDTTPGLPSPRLIC
               VAGQALYQADPAIFERHUKYISLIGKNGFSGVELCRDPLDGNTPVLVAVKQLQHS
               VPQQQDFQFEIOLIKALHSDFIVKRGVSGYFGFGRSLRMEYLPSCGLRDLQLQHR
               GLTDRLLLPWAOITCKMGYILGARCVHRDLAARNILVSEAHVKLADFGLAKLLPLG
               KDYVVVRPEQSPIFWYAPESLSDNIPFSROSDVMSFGVLYELFTVTCDKSCSPSFL
               RWMGPEREGPPLCRILLEAEGRRLLPPPTCTPEVQELMQLCWAPEPHDRPAPATLSL
               QLDPLEWRGRPG"
13..3309
   /gene="jak3"
   /product="protein tyrosine kinase"
   /standard_name="Janus kinase, mujak3"
   /evidence=experimental
3313..3723
   /gene="jak3"
3723
   /gene="jak3"
mat_peptide
3'UTR
polyA_site

```

```

BASE COUNT      711 a      1162 c      1079 g      771 t
ORIGIN
Alignment Scores:
Pred. No.:      0      Length:      3723
Score:      5854.00      Matches:      1098
Percent Similarity:      99.91%      Conservative:      0
Best Local Similarity:      99.91%      Mismatches:      1
Query Match:      99.90%      Indels:      0
DB:      10      Gaps:      0
US-09-397-967A-16 (1-1099) x MUSPTYKIN (1-3723)
QY      1 MetAlaProSerGluGluThrProLeuLeuLeuProGlnArgSerCysSerLeuSerSer 20
      |||||
Db      13 ATGGCACTTCCAAAGTGAGGAGACACCTCTGATCCCTCAGCGCTTTGCGACGCTCTCATCC 72
      |||||
QY      21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
      |||||
Db      73 TCAGAGGAGGAGGAGCTTGCAATGTGCTCTCTCCCGGGGACCTGGCGCTCCCGGCGA 132
      |||||
QY      41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValargalaAlaAla 60
      |||||
Db      133 TTGTCAATCTCTTTTGGGAGCTACTTTGGCTGAGGATTTATGTGTGGAGCTGCCAAGGCC 192
      |||||
QY      61 CysGlyLeuLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
      |||||
Db      193 TGTGCTATCTGCTGCTTTATCATTCGCTTTTCGCTTGCCACTGAGGACTTCTCTTGC 252
      |||||
QY      81 TrpPheProProSerHisPheCysIleGluAspValAspThrGlnValLeuValTyr 100
      |||||
Db      253 TGGTTTTCCCAAGCACACATCTTCTCATAGAGAGCGTGACACTCAAGTCTTTGGTCTAC 312
      |||||
QY      101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120
      |||||
Db      313 AGGCTACGCTTTTATTTCCCTGACTGTTTGGGCTGGAGACATGTCAACGCTTTGGGCTG 372
      |||||
QY      121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
      |||||
Db      373 CGCAAGATTTGACCACTGCCATCCTTGACTTACATGTTTATAGACATCTTTTGCTCAG 432
      |||||
QY      141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
      |||||
Db      433 CACCGCAGTGACCTGGTGAGTGGGCGCTCCCGTGGGCTTAGCATGAAGGAGCAGGGA 492
      |||||
QY      161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGluMetAlaArgGluGlnAlaGlnArg 180
      |||||
Db      493 GAGTTCTGAGCTTGCGCTGGGCTGGGCTTGAGCTTGAGCTTGAGCTTGAGCTTGAGCTTGAG 552
      |||||
QY      181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
      |||||
Db      553 CCAGGAGAGCTGCTGAAGAGGTCAGTTACAAAGCTGTCTGCCCGCCAGCCTGCGCGAT 612
      |||||
QY      201 ValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAla 220
      |||||
Db      613 GTGATCCAGGCGCAGAACTTCGTGACACGACGAGCGCATCCGCGAGGACCGGTGCTTGGCG 672
      |||||
QY      221 LeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeu 240
      |||||
Db      673 CTGCTCCGCTGGTGGCTCGCTCGCCAGGCGGACCGCTCATGGCCAAAGTATATTTCTG 732
      |||||
QY      241 AspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGly 260
      |||||
Db      733 GACCTGGAGCGGCTACATCCAGCGGCCACCCACGAGACCTTTCCTGTGTGGGCTCCCGGGC 792
      |||||
QY      261 AlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrpSerSer 280
      |||||
Db      793 GCCCAGGAGGAGCGGGGCTTCTGCGTGTGGCGAGGAGCAACCGCATCCCTTGGAGTCG 852
      |||||
QY      281 AsnAspGluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsn 300
      |||||
Db      853 AACGACGAGCTTTTCCAGACCTTCTGTGACTTTTCCCGAATATCGTGGATGTGACATCAAT 912
      |||||

```

QY 301 GlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAsp 320
DB 913 CAGGCCCCACAGGTGTGGTCCCGCAGGGAGACCCGGCTGGTCACTGTCAACAGGATGGAC 972
QY 321 GlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeu 340
DB 973 GGCCACATCCTTGGAAAGCGAGTTTCGGGGCTGGCTGAGGGCTGTCTTTCGTGGCCCTC 1032
QY 341 ValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysIleGluValAla 360
DB 1033 GTGGATGGGTACTTCGGCTGTATCTGGCACTCCAGGCATTATTTCTGCAAGGAGTGGCG 1092
QY 361 ProProArgLeuLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPhe 380
DB 1093 CCGCCAAAGGCTGCTGGAGAGGAGCGGACGTGTGCCATGAGCCCATCAGCTTAGACTTT 1152
QY 381 AlalIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSer 400
DB 1153 GCCATCCCAAGCTGAAGGCCGCTGGCTCCCTCCAGGCACCTATATTTCTCCGCGCAGC 1212
QY 401 ProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAsp 420
DB 1213 CCGCAGGACTATGACAGCTTTCCTTACCGCTCGCTCCAGACTCCTCTTGGCCCCGAC 1272
QY 421 TyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSer 440
DB 1273 TACAAGGGCTGCCTCATCCGCCAGACCCCGCGGGCTTTCTCCCTGGTTGGCCTCAGC 1332
QY 441 GlnProHisArgSerLeuArgGluLeuLeuAlaCysTyrPheAsnSerGlyLeuArgVal 460
DB 1333 CAGCCCCACAGAAGCTCGGGAGCTGCTTGCAGGCTGTCTGGAATTTCTGGGCTGGAGTA 1392
QY 461 AspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGluLysSerAsn 480
DB 1393 GACGGTGTCTGCCTGTACTTAACATCTCTGCTCGCTCCAGACCCCAAGAAAGTCCAAAT 1452
QY 481 LeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCys 500
DB 1453 TTGATCGTGGTGGGAAGGGCTGCAACCCCGCGCTGCCCTGGCTGTCCCGTCTCTGC 1512
QY 501 CysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluThrHisGlu 520
DB 1513 TGTGGCTGTACACAGCTGAGCTTCCACACAAATCCAAACGGAAGCTGAGTGGCAGCAG 1572
QY 521 AsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValValAsp 540
DB 1573 AACCTGGGTACCGTTCCTTTACCAAGATCTTCCGTGGCAGCAGCGGAGTCTGGAT 1632
QY 541 GlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCys 560
DB 1633 GGTGAGACACATGACTCGGAAGTCTCTCTGAAGTTCATGGACTCCAGACATCGGAACCTGC 1692
QY 561 MetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuVal 580
DB 1693 ATGGAGTCTTTCTTGGAAAGCCGACAGCTTGTATGAGCCCAAGTATCTACCCGACCTGGT 1752
QY 581 LeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTyrLeu 600
DB 1753 TTACTGACGGGCTGTGATGGCTGGAGACAGCATCATGGTGGCAGAAATTTGTGTATCTA 1812
QY 601 GlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrLysLeu 620
DB 1813 GGAGCAATTGACATGACTCTGGCAAGCGTGGCCACTTGGTGTGACAGCAGCTGGAAACTG 1872
QY 621 GlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHis 640
DB 1873 CAGGTGACCAAGCAGCTGGCATATGCCCTTAACCTACTTGGAGCAAAAGGCTTCTCTCAC 1932
QY 641 GlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGlyAspGlyAsnProPro 660
DB 1933 GGCAACGCTCAGCACGGAAGGTGCTCTGGCTGTGAGGGGGTGTATGGGAATCCACT 1992
QY 661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680

DB 1993 TTCATTAAAGCTGAGTGATCCTGGTGTCACTGCCATCTGTCTGAGCCTGGAATGTCTACC 2052
QY 681 AspArgIleProTyrValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 700
DB 2053 GACAGAATACCTGGTGGCCGCCCAATGTCTCCAGGAGGCTCAGACACTCTGCTTGGAG 2112
QY 701 AlaAspLysTyrGlyPheGlyAlaThrThrTyrGluValPheGlnArgGlyProAlaHis 720
DB 2113 GCTGCAAGTGGGCTTTGGAGCCACCGTGGAGGTGTTCCAAACGGGGAACCGGCCAC 2172
QY 721 IleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlnGlnLeuPro 740
DB 2173 ATCACCCTCGCTGGAGCCGCCCAAAAGCTGAAGTTCTATGAGACCGAGGACAGCTGCC 2232
QY 741 AlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760
DB 2233 GCTCTCAAATGACAGAACTGGCGGACTTATCACACAGTGCATGGCGTATGATCTGCGC 2292
QY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780
DB 2293 CGCGCCCCCTCTTCGAGCTATCTCAGAGACCTCAACGGCTCATTTACATCAGATTAC 2352
QY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAla 800
DB 2353 GAGCTCCTCTCAGACCCACACCTGGCATCCCGAGTCCCTCGAGATGAGCTGTGCGTAGCT 2412
QY 801 GlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLysTyr 820
DB 2413 GGCGCCAGCTTATGCTGTCAGACCCCGCATATTCGAGGAGAGACACCTTAAGTAC 2472
QY 821 IleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeu 840
DB 2473 ATCTCTTGTGGGCAAGGGCAACTTTGGCAGCGTGGAGCTGTGCCGTATGACCCCTG 2532
QY 841 GlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAspGln 860
DB 2533 GGGACAATACGGGACCCCTGGTGGCAGTGAACAGCTACAGCAGAGGTGCCACACAG 2592
QY 861 GlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVal 880
DB 2593 CAGAGGAGCTTCCAGCGGGAGATTGAGTCTTAAGGCTCTGCACAGCAGCTTCATCGTC 2652
QY 881 LysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTyr 900
DB 2653 AAGTACCGGGGAGTCAGCTATGGGCGCAGGTGCCAGAGCCCTGGCTTGGTATGAGTAC 2712
QY 901 LeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGlyLeuHisThrAspArg 920
DB 2713 TTGCCACGGCTGCCCTCGAGACTTGTGTCAGCGCCATCGCGGCTTGCACACCGACCGC 2772
QY 921 LeuLeuLeuPheAlaTyrGlnIleCysLysGlyMetGluTyrLeuGlyAlaArgArgCys 940
DB 2773 CTACTGTCTTTCGTGGCAGATCTGCAAGGGCATGGAGTACCTGGGTGCGCGCGCTGC 2832
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisValLysIle 960
DB 2833 GTACACCGTGACCTGGCTGCGCGCAACATCTTGTGGAGAGCGAGGCTCATGTGAAGATC 2892
QY 961 AlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArgGlu 980
DB 2893 CGGACTTCGGCTTCGTAACTGCTGCCCTCGGAAAGGACTACTAGCTGGTGGCGGAG 2952
QY 981 ProGlyGlnSerProIlePheThrTyrAlaProGluSerLeuSerAspAsnIlePheSer 1000
DB 2953 CCGGCAAAAGCCCATCTTTTGTGTATGCCCGGAGTCCCTATCTGACAAACATCTTCTCC 3012
QY 1001 ArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluLeuPheThrTyrCysAsp 1020
DB 3013 CGCCAATCTGACGTGTGGAGTTTCGGAGTGTGTGTACGAGCTCTTCCACCTACTGCGAC 3072
QY 1021 LysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGluGlyPro 1040

280 rAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIle 299
911 CGGGACCAGGAGCTTTTCAGACCTTCTGTGACCTTCGGAAATCGTGGATGTCAGCAT 970
299 eAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMe 319
971 CAAGCAGGCCCCACGCTGTGGGTCCGGCAGGGGAGCACCAGGCTGTCACCTGTCCACAGGAT 1030
319 tAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAl 339
1031 GGACGCCACATCTGGAGCGGAGTTTCGGGGCTGCTGAGGCGCTGCTTTTCGTGGC 1090
339 aLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluVa 359
1091 CCTCGTGATGGTACTTCCGCTGATCTGGACTCCAGGCATTTATTTCTCAAGAGAGT 1150
359 lAlaProProArgLeuLeuGluAlaAspValCysHisGlyProIleThrLeuAs 379
1151 GGCGCCGCCACGGCTGCTGGAGGAGGAGCGGAGCTGTGCCATGGACCCATCACGTTAGA 1210
379 pPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuAtqAr 399
1211 CTTTGGCATCCACAGCTGAGGCGCTGGCTCCCTCCAGCACCTATATTTCTCCGCCG 1270
399 gSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyPr 419
1271 CAGCCCGCAGGACTATGACAGCTTCTTCTTACCGCTGCGTCAGACTCTCTTTGGCCC 1330
419 oAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLe 439
1331 CGACTACAAGGGCTGCTCATCTCCGCCAGGACCCAGCGGGCTTTCTCCCTGGTGGGCT 1390
439 uSerGlnProHisArgSerLeuArgGluLeuAlaAlaCysTrpAsnSerGlyLeuAr 459
1391 CAGCCAGCCCCACAGAGCTTGGGGAGCTGCTGCAGCCTGCTGGAATCTTGGGCTGCG 1450
459 gValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGluLysSe 479
1451 AGTAGACGGTCTGCCCTGAACCTTAACATCTGCTGGCTCCAGACCCCAAGAAAGTC 1510
479 rAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSe 499
1511 CAATTTGATGCTGGTGGAGAGGCTGACCCCGCGCTGCTGCCCTGCTGCTGCTGCCGTC 1570
499 rCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHi 519
1571 CTGCTGTGGCTGACACAGCTGAGCTTCCACAAATTCACAGCAGCCTGGAGTGGCA 1630
519 sGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValVa 539
1631 CGAGAACCTGGGTCACTGTTTACCAAGATCTTCCGTGGCGCAGCGGGAGGTCGT 1690
539 lAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAs 559
1691 GGATGGTGAGACACATGACTCGGAAGTCTCTCGAAGTCAATGACTCCAGACATCGGAA 1750
559 nCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLe 579
1751 CTGCATCGAGTCTTTCTGGAGCGCGAGCTTGTATGAGCCAAAGTATCTTACCCGCACT 1810
579 uValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTy 599
1811 GGTGTTACTGACGCGCTCTGCATGGCTGGAGACAGCATCATGTGTGGAGAAATTTGTGA 1870
599 rLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLy 619
1871 TCTAGGAGCAATTCACATGATACCTGCGCAAGCGTGGCCACCTGGTGTGACGACCTGAA 1930
619 sLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuPr 639
1931 ACTGAGGTGACCAAGCAGCTGGCATATGCCCTTAACCTACTTGGAGGACAAAGGCCCTTC 1990
639 oHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGlyAspGlyAsnPr 659

1951 TCACGGCAACGCTCTCAGACCGAAGTGTCTCTGGCTCTGAGGGGGTGTATGGGAATCC 2050
659 oProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLe 679
2051 ACCTTTTCAATTAAGCTGAGTGATCTGTGTGTCTCAGTCCCACTGTGTCTGAGCTGGAAATGCT 2110
679 uThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLe 699
2111 CACCAGACAAATACCTCTGGGTGGCCCCGAAATGTCTCCAGAGGCTCAGACACTCTGCTT 2170
699 uGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAl 719
2171 GGAGGCTGACAAAGTGGGCTTTGGAGCCACCACTGGGAGGTTTCACGGGGGACCCGCG 2230
719 aHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyLe 739
2231 CCACATCACCTCGCTGGAGCCGCCAAAAGCTCAAGTTCTATGAGGACCAAGGACAGCT 2290
739 uProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPr 759
2291 GCCCGCTCTCAATGGACAGAACTGGCGGAGCTTATCACAGATGCATGGCGTATGATCC 2350
759 oGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs 779
2351 TGGCGGGCGCCCTCTCTCCAGCTATCTCCAGAGACTCAACCGGCTCATATACATCAGA 2410
779 pTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysVa 799
2411 TTACGAGCTCTCTCAGACCCACACCTGGCATCCGAGTCTCCGAGATGAGCTGTGC-- 2468
799 lAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLy 819
2469 -GGTGGCGCCCGAGCTCTATGCTCTCCAGGACCCCGCCATATTCAGGAGAGACACCTTAA 2527
819 sTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspPr 839
2528 GTACATCTTTGCTGGCAGGAGCACTTTGGCAGCGTGGAGGTGTGCCCTATGACCC 2587
839 oLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAs 859
2588 CCTGGGGACAATACGGGACCCCTGTGGCAGTGAACAGACTACAGACACAGCGGGCCAGA 2647
859 pGlnGluArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIl 879
2648 CCACAGAGGAGACTTCCAGCGGGAGATTCCAGATCCTTAAGGCTCTGCACAGCGACTTCA 2707
879 eValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGl 899
2708 CGTCAAGTACCGGGAGTCAAGTATGGCCAGGTCGCCAGAGCCTGGGTGGTATGGA 2767
899 uTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly---LeuHisTh 918
2768 GTACTCTCCACGCGCTGCTGCGAGACTTCTCAGCGCCATCTCAGCGGCTGCAAC 2827
918 rAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetClnTyrLeuGlyAlaAr 938
2828 CGACCGCTACTGCTGTCTTGGCAGATCTGCAAGGGATGAGAGTACCTGGTGGCGCG 2887
938 gArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisVa 958
2888 CCCTGGTGATACCGTGAACCTGGCTGGCGCGCAACATCTTGGTGAGAGCGAGGCTCATGT 2947
958 lLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVa 978
2948 GAAGATCGCGGACTTCCGCTCGCTAAGCTGTGCCCTGGCCCTGGGAAAGGACTACTACGTGGT 3007
978 lArgGluProGlyClnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnIl 998
3008 CCGCAGCCTTGGCCCAAGGCCCACTCTTTGGTATGCCCGGAGTCCCTATCTGTACACAT 3067
998 ePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTy 1018

Db 3068 CTTCTCCCGCAATCTGACGCTGCGAGCTTCGAGTGGTGTGTACGAGCTCTTACACCTA 3127
 Qy 1018 rCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgG1 1038
 Db 3128 CTGCGACAGAGCTGACGCCATCGCTGAGTTCCTGCGCATGATGGGCGCTGAGCGTGA 3187
 Qy 1038 uGlyProProLeuCysArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProProPr 1058
 Db 3188 AGACACCCCGCTCTCCGCGCTCTGAGAGTGTGCGAGAGGCCGAGCGCTCCACACACC 3247
 Qy 1058 oProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTTPAlaProGluProHi 1078
 Db 3248 TCCACCTCCCGCACCGAGGTTGAGAGCTCATGAGCTGTGCTGGCGCGCCAGCCGCA 3307
 Qy 1078 sAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTTPArgGlyArgPr 1098
 Db 3308 CCACCGCGCAGGCTTCGGCACCTGAGCCCGCAGCTGGAGCGGCTGTGGCGTGGAGAGCC 3367
 Qy 1098 oGly 1099
 Db 3368 CGGA 3371

RESULT 3

RATJAK3 3778 bp mRNA linear ROD 10-JAN-2003
 LOCUS Rattus norvegicus JAK3 mRNA for protein-tyrosine kinase, complete

DEFINITION
 ACCESSION D28508
 VERSION D28508.1 GI:485811
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1
 AUTHORS Takahashi, T. and Shirasawa, T.
 TITLE Molecular cloning of rat JAK3, a novel member of the JAK family of protein tyrosine kinases
 JOURNAL FEBS Lett. 342 (2), 124-128 (1994)
 MEDLINE 94192816
 PUBMED 8143863

REFERENCE 2 (bases 1 to 3778)
 AUTHORS Shirasawa, T.
 TITLE Direct Submission
 JOURNAL Submitted (10-FEB-1994) Takuji Shirasawa, Tokyo Metropolitan Institute of Gerontology, Department of Molecular Pathology; 35-2 Sakae-cho, Itabashi-ku, Tokyo 173, Japan
 (Tel:81-3-3964-3241(ex.3034), Fax:81-3-3579-4776)

FEATURES

source
 1..3778
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="JAK3"
 /tissue_type="spleen"
 /dev_stage="adult"
 1..3778
 /gene="JAK3"
 61..3363
 /gene="JAK3"
 /evidence="experimental"
 /product="protein-tyrosine kinase"
 /protein_id="BA005868.1"
 /db_xref="GI:485812"

gene
 CDS
 1..3778
 /translation="MAPSSETELISQSSLSSEAGALHLLPRGPPOPLRFS
 FGYLARDLCVRAKACGILPVHSLPALATEDLSWFPSPSHIFSIEDVDTQVLVRL
 RYFPGWGLETCRFGKLDLTSAILDVHLEHFAOHRSLVSGRLPVGLSLKDOG
 EFLVLDLDAQARKQAPRGELLKSVKACLPSSLRLDIOGSPVTRRRIRTVV
 QALAPCSLSRFPYALWAKYILDLERLHPAATTESFLVGLPQAGEFGCLRVTDNGI
 AWSKQDELFTQCFDPEIIVDSIKQAPRVGPAGEHRLVTITRMDGHLEAFPGLPF
 ALSFVALVDGYFRILICDSRHFFCKEVPAPPELLEBEEAELCHGPIITLDFAIHKLKAAGSL

BASE COUNT 729 a 1172 c 1060 g 817 t

ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 3778
 Score: 5380.50 Matches: 1018
 Percent Similarity: 94.64% Conservative: 24
 Best Local Similarity: 92.46% Mismatches: 56
 Query Match: 91.82% Indels: 3
 DB: 10 Gaps: 3

US-09-397-967A-16 (1-1099) x RATJAK3 (1-3778)

Qy 1 MetAlaProSerGluGluThrProLeuLeuProGlnArgSerCysSerLeuSerSer 20
 Db 61 ATGGCACCTCCAAAGCGAGGAGCGCTCTCACTCTCAGGGCTCTTAGCCTCTCATCC 120
 Qy 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
 Db 121 TCAGAAAGCAGGAGCCCTGCATGTCTCTCTCCCGGGGACCTGGCGCTCCCGACGGA 180
 Qy 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
 Db 181 CTGTCACTCTCTTTGGGACTACTTTGGCTGAGGATCTGTGTGTGGGGCTGCCAGGCC 240
 Qy 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
 Db 241 TGTGGCATCTGCCCGCTTTATCATTCGCTCTTCGCTCGCCACCGAGGACTTGTCTTTC 300
 Qy 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
 Db 301 TGGTTCCCCCAAGCACATCTTCTCCATAGAGGAGCTGGACACTCAAGTCTTTGGTCTAC 360
 Qy 101 ArgLeuArgPheTyrPheProAspTTPheGlyLeuGluThrCysHisArgPheGlyLeu 120
 Db 361 AGGCTACGCTTTATTTTCTGTTGTTGGACTGGAGACCTGTACCCGCTTTGGGTTA 420
 Qy 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
 Db 421 CACAAGATTGGACAGTGCCATCTTTCAGTACATGTTTTAGAACACCTCTTTGGCCAG 480
 Qy 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
 Db 481 CACCGCAGTGACCTGTGTAGTGGCGCTCTCCAGTAGGCGCTTAGCGCTGAAGGACCAAGGA 540
 Qy 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
 Db 541 GAGTTCTGAGCCTAGCTGTGTGGACCTGGCCAGATGGCTGTGAGCGAGGCCACCGCG 600
 Qy 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
 Db 601 CCAGGAGAGCTGTGAAGTGTCTGAGTTACAAAGCCTGTCTGGCGCCACGCTGCGCGAC 660
 Qy 201 ValIleGlnGlnAsnPheValThrArgArgArgIleArgThrValValLeuAla 220
 Db 661 CTGATCCAGGCGCAGAGCTTCTGTGACGCGCAGGCGCATCCGCGAGCCGCGTGTGAGCGG 720
 Qy 221 LeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeu 240
 Db 721 CTGGCGCGGTGTAGTACCTTCCAGCGGACCCCTACGCACTCATCGGCCAAGTACATCCTA 780

QY	241	AspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGly	260
DB	781	GACCTTGAGCGGCTGCATCCAGCGCCACCACTGAGTCCTTCTCGTGGAGACTCCCGGGC	840
QY	261	AlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyVilProTrpSerSer	280
DB	841	GCCCAGGAGGAGCCGGGTGCTCGTGTGACAGGGGATATGCGATCGCTGGAGTCC	900
QY	281	AsnAsp--GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIle	299
DB	901	AAGGACCAGGAGCTTTTCCAGACCTTCTGTGACTTCCCGAAATCGTGGATGTGAGCATC	960
QY	300	AsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMet	319
DB	961	AAGCAGGCTCCACCGTGGGTCCCGCAGGGAGCAGCGGTAGTCACTACACAGAGTG	1020
QY	320	AspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAla	339
DB	1021	GACGGTCACATCCTGGAAGCGGAGTCTCCGGGGCTGCCTGAGGGCGCTGTCTTCTGTGGCG	1080
QY	340	LeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluVal	359
DB	1081	CTCGTGATGGGTACTTTTCGCCTCATCTGCACCTCCAGGCAATTTCTTCTGCAAGAGGTG	1140
QY	360	AlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIleThrLeuAsp	379
DB	1141	GCACCGCACGGCTGCTGGAGAGGAGCGGAGCTGTGCCATGACCCCATCGTTAGAC	1200
QY	380	PheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArg	399
DB	1201	TTTGCCATCCACAAGCTGAAGGCTGCTGCCTCCCTCCGGGCTCTTACATCTCCGACGC	1260
QY	400	SerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyPro	419
DB	1261	AGCCCGCAGACTATGACAGCTTTCTTCTTACTGCCTGTGTCCAGACTCCTCTTGGCCCC	1320
QY	420	AspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeu	439
DB	1321	GACTACAAAGGCTGCTTCATCCGCCAAGACCCAGTGGGGCTTTCTCCCTGGTGTGCTC	1380
QY	440	SerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArg	459
DB	1381	AGCCAGCTGCACAGAACCCTACAGGAGCTGCTCACAGCTGCTGGCATTTCTGAGCTGCAG	1440
QY	460	ValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaProArgProLysGluLysSer	479
DB	1441	GTAGCGGCACTGCCTTGAACTCATCTTGTGTGCTGCCAGACCCAGAACCAAGGAAAGTCC	1500
QY	480	AsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSer	499
DB	1501	AATTTGATCTGTGTGTAGGGCGGCAACCCACGCTCTCCCGGGCACTCCCATCC	1560
QY	500	CysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTyrHis	519
DB	1561	TGCTGTGCGCTGACTAAGTTGAGCTTCCACAAATCCACAGCAGACGCTGGAGTGGCAC	1620
QY	520	GluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValVal	539
DB	1621	GAGAACTTGTGTACGGCTCTTTTACCAGATCTTCCATGGCCATAGCGGGAGGTTGTG	1680
QY	540	AspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsn	559
DB	1681	GATGGTGAGACATGACACAGAAGTCTCTCTGAAGGTTATGAGCTCCAGACACCGAAG	1740
QY	560	CysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeu	579
DB	1741	TGCATGGAGTCTTTCTGGAGCGCGAGCTTGATGAGCCAAAGTATCTTACCCGCACTTG	1800
QY	580	ValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTyr	599
DB	1801	GTGTTGTGATGGAGTCTCATGGCTGGAGACAGCATCATGTGTGAGGAAATTTGTGATC	1860

```

Db      2938  AAGATTGCTGACTTGGCTCGCTAAGCTGCTGCTCTTGGAAAGACTACTACGTGTC 2997
Qy      979   ArgGluProGlyGlnSerProIlePheThrTyAlaProGluSerLeuSerAspIle 998
Db      2998  CGCGTGGCGGGCCAAAGTCCCATCTTCTGTTAGCGCCCTGAGTCCCTATCTGACACATC 3057
Qy      999   PheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyGluLeuPheThrTy 1018
Db      3058  TTCTCCCGCAATCTGACGTGGAGCTTGGCGTGGTCTGTATGAGCTCTTCCACTAC 3117
Qy      1019  CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGlu 1038
Db      3118  ACGGACAAAGCTGACAGCCCATCCACTGATTCCTCGGTATGATTGGCGCTGAGCGTGAG 3177
Qy      1039  GlyProProLeuCysArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProPro 1058
Db      3178  GGATCCCGCTCTGCCACCTCTTGGAGCTGCTGGCAGAGCGCGCTCTCCACACCC 3237
Qy      1059  ProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGluProHis 1078
Db      3238  TCCACTGCTCTACTGAGGTTCAAGAGCTCATGACGCTGTGCTGGCCCAACCCACAA 3297
Qy      1079  AspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgPro 1098
Db      3298  GACCGGCGCGGCTTGACACCTTGAGCCCGGAGCCCGGCTGTGGCGCGGAGGCCCC 3357
Qy      1099  Gly 1099
Db      3358  GGA 3360

```

```

RESULT 4
MUSJAK3H
LOCUS   Mus musculus (clone 32DS-1) protein tyrosine kinase (JAK3) mRNA.
DEFINITION
ACCESSION L33768
VERSION L33768.1 GI:508533
KEYWORDS protein tyrosine kinase.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Rane, S.G. and Reddy, E.P.
TITLE JAK3: a novel JAK kinase associated with terminal differentiation
of hematopoietic cells
JOURNAL Oncogene 9 (8), 2415-2423 (1994)
MEDLINE 94309920
PUBMED 7518579
COMMENT Original source text: Mus musculus cDNA to mRNA.
FEATURES
source Location/Qualifiers
1. .4016
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="32DS-1"
/call_line="32Dc13"
/tissue_lib="lambda gt11"
7. .3906
/feature="protein tyrosine kinase"
/codon_start=1
/product="JAK3"
/protein_id="AAA21415.1"
/db_xref="GI:537607"
/translation="MAPPSSETPLIPQSCSLSSSEAGALHVLPRGPPQRLSPS
FGDYAEDLCVRAAKACILPVHVSUFALATEDFSCWFFPSHIFCIEDVQVLVYRL
RFYPPDFGLTCHRFGRKDLTSAIDLHLFHLFAQHSRLVSGRLPVGLSMKEG
EFLSLVDLDAQMARQOAPGELLKTVSKACILPFLSRDVIQQNPVTRIRRTTV
LALRVSPARTATAGQVLDLRLHPAATTFTRVGLPQAQEPGLRLVAGDNGLS
WSSDQGLGLSLRGVGRNRLVKPVTIRKVASPGRADCTTGCGGGLNKGVP
MELPQGLVGTTRLDNRGRKTEGSDMWISGHDTRFVPSLTSPPPHKWRWG
GRGCGAGSRVLPWLLSLFLFFFNFGARQFSYSSGCGTTFVRPGWPTQKACL
CLSSAVIKGRVPLRLYCLSLFQLPQFCDFPEIVDVSIKQPTCGSGREHRLVTTRM

```

```

DHLLEAFPPGLPEALSVALVDGYFRLICDSRHYFCKEVAAPRLLEEAEILCHGPIIT
LCASQGWGMAARTVLGWTLLSWPRLDFAIHKLAASLPGTIILRRSPQYDS
FLITACVDTPLGDKGLIPDSGASLNASASPTACSCGLLEFGLRVDGAAL
NLTSACAPRPEKSNLIVVRGCTPAPAGSPSCALTLQSFHTIPTSLGSHENLGH
GSFTKIFRRRRVVDGTHSEVLLKWDSEHNCMESFLAASLMSOVSYPHLVLL
HGVCMAGDSIMVFEVYLGADIMYLKRKGLHVSWSKLVQTKQLAYALNLEDKGLPH
GNVSARKVLLAREGGDGNPFIKLSDDPGVSPVLSLEMLTDRIFWAPCELOEATLCL
LEADKMGCAATWTFVSGGPAHITSLEPAKLKFYEDDQQLPALKWEIAGLITQMA
YKRRPFRATRLDLNGLITSYDLNLTGFLAVKOLSHGSCAWRPAICLPGRHIEER
HLKYSLLKGNFGSVELCRYPDLNLTGFLAVKOLSHGSCAWRPAICLPGRHIEER
SDFIYKRGVSGPGRQSLRLVMEVPSGLDFLQHPAALHTDLRLLLFAMQICKGM
EYLGAARRCHRDLAARNILVSEBAHVKAIDFGLAKLLPLCKDYIVVREPGSQPIFWA
PELSNDNFISQSDVMSFVLYELFTYCDKSCSPSAEFLRMWGPEREFPFLCRLLEL
LAEGRRLLPPPTCTPTPELMOQLCVAPSRITQPSAP"
misc_feature 2199..3024
/feature="JAK homology domain II (JH2)"
misc_feature 3076..3876
/feature="JAK homology domain I (JH1)"
BASE COUNT 742 a 1203 c 1235 g 836 t
ORIGIN
Alignment Scores:
Pred. No.: 3 05e-311 Length: 4016
Score: 5223.00 Matches: 1060
Percent Similarity: 80.59% Conservative: 3
Best Local Similarity: 80.36% Mismatches: 32
Query Match: 89.13% Indels: 235
Gaps: 6
US-09-397-967A-16 (1-1099) x MUSJAK3H (1-4016)
Qy 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
Db 7 ATGGCACTCCCAAGTCAGGAGACACCTCTGATCCCTCAGCGCTCTTGCAGCCTCTCATCC 66
Qy 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
Db 67 TCAGAGCGAGGAGCCCTGTCATGTGCTCTCTCCCGGGGACCTGGCGCTCCCGACGGA 126
Qy 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
Db 127 TTGTCATCTCTTTTGGGACACTCTGGCTGAGAGATTATGTGCGAGCTGCCAGGCC 186
Qy 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db 187 TGTGCGATCCTGCTGCTTTATCATCTGCTTTTGGCTCTGGCCACTGAGGACTTCTCTTGC 246
Qy 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db 247 TGGTTTCCCAAGCCACATCTCTGATAGAGAGCGTGGACACTCAAGTCTTGGTCTAC 306
Qy 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120
Db 307 AGGCTACGCTTTTATTTCCCTGACTGTTTGGCTGGAGCATGTACCCGCTTGGGCTG 366
Qy 121 ArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
Db 367 CGCAAGATTGACCACTGCTTGCATAGAGAGCGTGGACACTCAAGTCTTGGTCTAC 426
Qy 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGly 160
Db 427 CACCGCAGTACCTGGTGGCGCTCCCGTGGCGCTTACAGTGAAGGAGCAGGGA 486
Qy 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
Db 487 GAGTTCCTGAGCTGGCGCTGCTGGACTTGGCCAGATGGCTCTGAGCAGGCCAGCGC 546
Qy 181 ProGlyGluLeuLeuTyrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
Db 547 CCAGGAGAGCTGCTGAAGACGGTCAGTTACAAAGCCTGTCTGCCGCCAGCCTCGCGAT 606
Qy 201 ValIleGlnGlnAsnPheValThrArgArgArgIleArgThrValValLeuAla 220

```

607 GTGATCCAGGCCAGAACTTCGTGACACGACGCGCGATCCGACGAGACCGTGGTCTTGGCG 666
Db
221 LeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeu 240
Qy
667 CTGC--GCGTGTGTGCTCCAGGCCGACGCTACGCGTCATGTCACCAAGTATAT-CTG 723
Db
241 AspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGly 260
Qy
724 GACCTGGAGCGCTACATCCAGCGGCCACACCGAGACCTTCGCTGGGGCTCCCGGGC 783
Db
261 AlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrSerSer 280
Qy
784 GCCCAGGAGGAGCGCGGCTTCCTGCTGTGGCGGGGACAAACGGCATCTCCTGGAGCTCC 843
Db
281 AsnAsp----- 282
Qy
844 GGGACACGAGGCTGTGGGCTGGGCTTGAGAAAGCGGGGTCCGCGGGGAGCGGTGGCGG 903
Db
282 ----- 282
Qy
904 TTAGTAAGAAGCCCGTCCGAACATCAGGAAGTGGCTAGTCCCGGGCGGCTGATGTC 963
Db
282 ----- 282
Qy
964 ACTACGGGCGAGGGCGGGGTCAAATTTAAAGGTGGGTCCGGGATGGAGCTGCGCCAG 1023
Db
282 ----- 282
Qy
1024 GGACTGACCTGGGCGGTGACTAGAAAGGTGAGACTAGATAGGGGGCGGGGACGACAGAA 1083
Db
282 ----- 282
Qy
1084 GGGATAGTATGATTTGATCAGTGGGCATGACCCGACCGAGCCCTGTCTTCAGTCCCCCTA 1143
Db
282 ----- 282
Qy
1144 ACCTCGTCCCCCCCCCACAAGTGGAGGTGGGAAGGAGGAGGAGGTGTGCT 1203
Db
282 ----- 282
Qy
1204 GGGTCAAGGTCTGAATCCCTGGCTTCTGTCTCTCTTTCTTTTCTTTTAAATGGT 1263
Db
282 ----- 282
Qy
1264 TTTCCGAGACAGGGTTTCTCGTATAGCTCTGGTGTCTGGAACTCCTTTGTAAGACCA 1323
Db
282 ----- 282
Qy
1324 GGCTGGCCTCGAACTCAGAAATCCGCTCTGCTCTCGAGTGTGTGATTAAGGG 1383
Db
283 ----- 282
Qy
1384 CGTGTGCCACTACCGCGCTACTGCTGTCTTTCTTCCTCAGCTTTTCCAGACCTTCTGT 1443
Db
290 AspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyProAlaGly 309
Qy
1444 GACTTTCGGAAATCGTGGATGTGAGCATCAAGCAGC--CACCTGTGGTCCGCGAGG 1501
Db
310 GluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPhePro 329
Qy
1502 AG-CACCGGTGTGTCACTGTCCAGGATGACGCCACATCTCTGGAAGCGAGTTTCG 1560
Db
330 GlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIleCys 349
Qy
1561 GGGTGTCTGAGGGCGTGTCTTTCGTGGCCCTCGTGGATGGGTACTTCCGCTCATCTGC 1620
Db
350 AspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGluAla 369
Qy
1621 GACTCCAGGCATATTATCTGCAAGAGAGGTGGCGGCCACCGCTGCTGGAGAGAGGCG 1680
Db
370 AspValCysHisGlyProIleThr----- 377
Qy
1681 GAGCTGTGCCATGACCCATCACGTTAGGAGCTAGTTGGGGTCAACAGTGGGGATGGGA 1740
Db

378 -----LeuAspPhe 380
Qy
1741 TGGCAGCGAGGACTGTCTTGGGTGGACCTGGCTGTCTTGGCCCTAGGTAGACTTT 1800
Db
381 AlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgSer 400
Qy
1801 GCCATCCACAAGCTGAAGCGCGCTCCCTCCAGGACCTATATTTCTCCGCGCAGC 1860
Db
401 ProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAsp 420
Qy
1861 CCGCAGGACTATGACAGCTTTCTTTACCGCTCGCTCCAGACTCTCTTTGGCCCCGAC 1920
Db
421 TyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSer 440
Qy
1921 TACAAGGCTCCCTCATCCGACAGACCCAGCGGGCTTTCTCCCTGGTTGC-CTCAGC 1979
Db
441 GlnProHisArgSerLeuArgGluLeuAlaAlaCysTrpAsnSerGlyLeuArgVal 460
Qy
1980 CAGCCCCACAGAAGCTGCGGAGCTGCTTGACG-TGCTGGAATTC-GGCTGGAGTA 2037
Db
461 AspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGluLysSerAsn 480
Qy
2038 GACGCTGTGCTGCTGAACCTAACATCTCTGCTGCGCTCCAGACCAAGAAAAGTCCAAT 2097
Db
481 LeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCys 500
Qy
2098 TTGATCGTGTGGAGAGGGCTGACCCCGCGCTGCCCCCTGCTGCTCCCGTCTGCG 2157
Db
501 CysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGlu 520
Qy
2158 TGTGCGCTGACACAGCTGAGCTTCCACACAATTTCAACGGACAGCCTG---GGACACGAG 2214
Db
521 AsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValValAsp 540
Qy
2215 AACCTGGGTCACTGCTTTTACCAAGATCTTCGTCGCGCAGCGCGGAGGTGTGGAT 2274
Db
541 GlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCys 560
Qy
2275 GGTGAGACACATGACTCGGAAGTCTCTGAAAGTCAATGAGCTCCAGACATCGGAACATGC 2334
Db
561 MetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuVal 580
Qy
2335 ATGAGTCTTTTCTGGAAGCGCGCAAGCTTGATGAGCCAGATATCTACCCGACCTGGT 2394
Db
581 LeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTyrIleu 600
Qy
2395 TTACTGCACGCGCTCTGCTGCTGGAGACAGCATCATGCTGCGAGGAATTTGTGTATCTA 2454
Db
601 GlyAlaIleAspMetTyrLeuArgLysArgGlyHisIleuValSerAlaSerTrpLysIleu 620
Qy
2455 GGACAAATTCATGTACCTGCGCAAGCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 2514
Db
621 GlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHis 640
Qy
2515 CAGGTGACCAAGCAGCTGGCATATGCCCTTAACTACTTGGAGGACAAAGGCTTCTCTAC 2574
Db
641 GlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGlyAspGlyAsnProPro 660
Qy
2575 GGCAACGTCTCAGCACGAGAAAGTCTCTGCTGCTGAGGGGGGTGATGGAAATCCACCT 2634
Db
661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuMetLeuThr 680
Qy
2635 TTCAATTAAGCTGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2694
Db
681 AspArgIleProTrpValAlaProGluCysLeuGlnAlaGlnThrLeuCysLeuGlu 700
Qy
2695 GACAGAATACCTGGGTGGCCCCCGAATGCTCCAGAGGCTCAGACACTCTGCTTGGAG 2754
Db
701 AlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHis 720
Qy
2755 GCTGACAAAGTGGGGCTTTGGAGCCACACCTGGGAGGTGTTTCAGCGGGGAGCCCGCCAC 2814
Db

QY 721 IleThrSerLeuGluProAlaIylsLysLeuLysPheTyrGluAspGlnGlyGlnLeuPro 740
Db 2815 ATCACTCTGCTGGAGCCGCCAAAGAGCTGAGTTCTATGAGGACGAGGACGCTGCC 2874
QY 741 AlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760
Db 2875 GCTCTCAATGGACAGACTGGCGGAGCTTATCACACAGTGATGCGTATGATCTGGC 2934
QY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780
Db 2935 CGCGCGCCCTCTTCGAGCTATCTCAGAGACTCAACGGCTCATATCATCAGATTAC 2994
QY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAla 800
Db 2995 GAGCTCTCTCA-GACCCACACTTGGCATCCGAGTCTTCGAGATGAGCTGTGGT---- 3049
QY 801 GlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGlnArgHisLeuLysTyr 820
Db 3050 GCGCGCCAGCTCTATGCTGCCAGGACCCCGCATATC-GAGGAGAGACACTTAAGTAC 3108
QY 821 IleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeu 840
Db 3109 ATCTCTTTTGTGGCAAGGCAACTTTGGCAGCTGGAGCTGTGCCGTATGACCCCTTG 3168
QY 841 GlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAspGln 860
Db 3169 ---GACAATACGGGACCCCTGTGGCAGTGAACACTACAGCAGCGGCGCAGACCAG 3225
QY 861 GlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVal 880
Db 3226 CAGAGGGACTTCCAGCGGAGATTAGATCTTAAAGCTCTGCACAGCGACTTATCATGTC 3285
QY 881 LysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgIleValMetGluTyr 900
Db 3286 AAGTACCGGGAGCTAGCTATGGCCAGGTCGCGCAGAGCTCGCGTGTGTGGAGTAC 3345
QY 901 LeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg--GlyLeuHisThrAsp 919
Db 3346 CTGCCCCAGGGCTGCTTGGAGACTTCTTGACGCGGCATCGCGGCCCTTGACACCGAC 3405
QY 920 ArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArgArg 939
Db 3406 CGCCTACTGCTGCTTGGCAGATCTGCAAGGCGATGGAGTACCTGGGTGGCGCGCCG 3465
QY 940 CysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisValLys 959
Db 3466 TGGGTACACCGTACCTGGCTGGCGCAACATCTTGGTGAGAGCGAGGCTCATGTGAAG 3525
QY 960 IleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArg 979
Db 3526 ATCGCGGACTTCGGCTCGCTAGCTGCTGCCCTGGGAAAGGACTACTAGTGTGGC 3585
QY 980 GluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnIlePhe 999
Db 3586 GAGCCTGGCCAAAGCCCATCTTTGTATGCTCCCGAGTCCCTATCTGCAACATCTTC 3645
QY 1000 SerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyrCys 1019
Db 3646 TCCCGCAATCTCAGGTGGAGCTCGGAGTGGTGTGTGACGAGCTTCTTCACTACTGC 3705
QY 1020 AspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGluGly 1039
Db 3706 GACAAGAGCTGCAGCCATCCGTGAGTTCTCGCATGATGGGCTGAGCGTGAAGGA 3765
QY 1040 ProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgArgLeuProProProPro 1059
Db 3766 CCCCCTCTGCGCCCTCTGAGCTGCTGGCAGAGCGGCGAGCTCCACACCTCCC 3825
QY 1060 ThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGluProHisAsp 1079
Db 3826 ACTGCCCCACGAGGTTAGAGCTCATGCACTGTGCGTGGCGGCCAG-CGCGACGAC 3884
QY 1080 ArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgPro 1098

Db 3885 CGGCACGCTTCGGCACCTTGAGCCCGACGCTGGACGCGCTGTGGGTGGAAGACCC 3941
RESULT 5
AX203081
LOCUS
DEFINITION
ACCESSION
VERSION
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 4064
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BAGE COUNT 746 a 1292 c 1168 g 858 t
ORIGIN
Alignment Scores:
Pred. No.: 3,82e-278 Length: 4064
Score: 4684.00 Matches: 905
Percent Similarity: 87.61% Conservative: 64
Best Local Similarity: 81.83% Mismatches: 129
Query Match: 79.93% Indels: 9
DB: 6 Gaps: 5
US-09-397-967A-16 (1-1099) x AX203081 (1-4064)
QY 1 MetAlaProProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
Db 96 ATGGCACCCTCCAAAGTGAAGAGCGCCCTGATCCCTCAGGTTCATGACGCTTGTGTC 155
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
Db 156 ACGAGGCTGTGGCTGCTGATGTGCTGCCGCTCGGCGCCCGGGCGCCCGCCAGCGC 215
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
Db 216 CTATCTTCTCTTTCGGGACCATCTTGGCTGAGGACCTGTGCTGAGGCTGCGCTGCAAGGCC 275
QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db 276 AGCGGATCTCGCTGTGTACCATCTCTTGTCTTGCCACGAGGACCTGTCTGTC 335
QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db 336 TGGTTCCTCCCGAGGACCATCTTCTCCGTGGAGATGCCAGACACCAAGTCTGTGTATC 395
QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
Db 396 AGGATTCTGCTTTTACTTCTCCCAATTTGGTGTGGGCTCGAGAAGTGCACCGCTTCGGGCTA 455
QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
Db 456 CGCAAGGATTTGCCAGTGTCTATCTTGTACCTGCCAGTCTCGAGACCTCTTTGCCAG 515
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
Db 516 CACCGCAGTGACCTGTGTGAGTGGCGCCTCCCGTGGCGCTCAGTCTCAAGGAGCAGGGT 575
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
Db 576 GAGTGTCTCAGCTTGGCCGCTGTGGACCTTGCCCGGATGGCGGAGAGCGAGCCCGCGG 635


```

Db      2792  GCTGTCATGAGTAGTACCTGCCAGCGGCTGCTTGGCGGACTTCTCGAGCGGACCGCGC 2851
QY      915  y---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGlnTy 934
Db      2852  GCGCTCGATGCGAGCGGCTCTCTTCTATTCTCGCGAGATCTGCAAGGGCATGGAGTA 2911
QY      934  rLeuGlyAlaArgAtgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSe 954
Db      2912  CTTGGGCTCCCGCGCTGGTGCGACCGACCTGGCGCGCCGAAACATCTCTGGTGGAGAG 2971
QY      954  rGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAs 974
Db      2972  CGAGGCACACGTCACAGATCGTCACTTCGCGCTAGCTAAGCTGCTCGCGCTTGACAAAG 3031
QY      974  pTyrTyrValValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLe 994
Db      3032  CTACTACGTGGTCCGCGAGCGACCGACGCCATTTCTTGATATGCCCGCGAAATCCCT 3091
QY      994  uSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGl 1014
Db      3092  CTCGGACACATCTTCTCGCCAGTCAGACGCTTCGGGCTCGTCTGTACGA 3151
QY      1014  uLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGl 1034
Db      3152  GCTCTTCACTACTCGGACAAAGCTGCAGCCCTCGCGCGAGTTCCTGCGGATGATGGG 3211
QY      1034  yProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgAr 1054
Db      3212  ATGTAGCGGGATGTCCCGCCCTCTCGCCCTCTTGGAACTCTCGAGGAGGCGGACAG 3271
QY      1054  gLeuProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAl 1074
Db      3272  GCTGCGCGCCCTCTCGCTCGCTGCTGAGTTCACGAGCTCATGAGCTGTGCTGGGC 3331
QY      1074  aProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTr 1094
Db      3332  CCCTAGCCACAGAGCGGCGCATATTACGGCCCTTGGCGCCCTGGCGCCAGTGGACATGCTGTG 3391
QY      1094  pArgGlyArgProGly 1099
Db      3392  GAGCGGAAGCGGGGG 3407

RESULT 6
HSU09607
LOCUS   HSU09607          4064 bp      mRNA          linear      PRI 09-JUL-1994
DEFINITION   Human JAK family protein tyrosine kinase (JAK3) mRNA, complete cds.
ACCESSION   U09607
VERSION     U09607.1
KEYWORDS    GI:508730
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1124)
AUTHORS    Kawamura,M., McVicar,D.W., Johnston,J.A., Blake,T.B., Chen,Y.,
            Lal,B.K., Lloyd,A.R., Kelvin,D.J., Staples,J.B., Ortaldo,J.R. and
            O'Shea,J.
            Molecular cloning of L-JAK, a Janus family protein-tyrosine kinase
            expressed in natural killer cells and activated leukocytes
            Proc. Natl. Acad. Sci. U.S.A. 91 (14), 6374-6378 (1994)
JOURNAL    MEDLINE 94294384
PUBMED     8022790
REFERENCE   2 (bases 1 to 4064)
AUTHORS    O'Shea,J.
            Direct Submission
            Submitted (11-MAY-1994) John O'Shea, Leukocyte Cell Biology Section
            LEI BRMP, National Cancer Institute FCRDC, Bldg 560 Rm 3146 FCRDC,
            Frederick, MD 21702, USA
            Location/Qualifiers
            1..4064
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
FEATURES             source

```

```

/cell_line="YT, HUT-78"
/cell_type="natural killer cell, PHA activated T cells"
1..4064
/gene="JAK3"
96..3470
/gene="JAK3"
/codon_start=1
/product="JAK family protein tyrosine kinase"
/protein_id="AAA19626.1"
/db_xref="GI:508731"
/translation="MAPSEETPLIPQSCSLSTEAGALHVLIPARPGPQRLSFS
FDHLEUADVOAKASGILPVHSLFALATEDLSWFPFPHSHFVEDASTQVLLYRI
RYFFNWFLEKCHRFGRKDLASAILDLVLEHFAQHRSDLSVGRPLVSLKXEQ
ECLSLAVDLARMAREQAQPGELTKVSKACLPPSLRDLIQGLSFVTRRAIRRTVR
RALPRVAACQADRHSIMAKYIMDLERLDPAGAAETPHVGLPGALGHDGILRLVAGD
GGIANTQGRQEVLPFCDFEIVDISIKOAPRVGAGEHRLVTVTDTNQIILEAEPFG
LEASLFFVALVGVYFRLTDSOHCFCKEYVAPRLLIEVAEQCHGPTLLDPAFNKLTG
GRPSYVLRRSPQDFRSLFVVCVONPLGPDYKGLIRRSPTGFLVGLSRPHSSL
RELLATCMGGJHVGVAVLTSCCIPRPEKSNLIVQGRHSPTSLVQSQYQL
SMTFHKIPADSLWEHNIHGHSFTKIYRCRHEVVDGEARKTEVLLKVMDDKHKCM
BSFLBAASLMVSQSVYRHLVLLHGVCMAGSTMVQEFVHLGADMLRKRGHLVPASWK
LOVVKOLAVALNVEDKGLPHGNVARKVLLAREGADGSPFIKISDPGVSPAVLSLE
MLTDRIPWVAPCLBEAOTLSLEADKMGFGATWVEVSGVTWPIISALDPAKLOFYED
ROOLPAKWTEIALLIQQCMAYEPVORPSFRAVIRDLNLSLSDRYELLSDPTGALAP
RDLGNGAQLYACQDPTIPEERHLKYISQLGKNGFSGVSELCTFDPHAHRTGALVVKQ
LOHSGPDQQRDFQREITQILKLSHDFIVKRGVSGVGRPELRLMEYLPSCGLRDL
QRHARLDASRLLYSSQICKMGVYLRRCVVRDLAARNILVSEAHKVIADFLGAK
LAPLDKDYVVRPGVSPFVWAPESLSDNIFSRQSDVMSFGVLYELFTYCDKSCSP
SAELFMMCCERDVPAICRLLELLEGORLPAPPACPAEVHELMKLCWAPSPQDRFSF
SALGPOLDMLWSGRCETHAFTHAEPGKHHSLSFS"
BASE COUNT 746 a 1292 c 1168 g 858 t
ORIGIN
Alignment Scores:
Pred. No.: 3,82e-278 Length: 4064
Score: 4684.00 Matches: 905
Percent Similarity: 87.61% Conservative: 64
Best Local Similarity: 81.83% Mismatches: 129
Query Match: 79.93% Indels: 9
DB: Gaps: 5
US-09-397-967a-16 (1-1099) x HSU09607 (1-4064)
QY 1 MetAlaProProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
Db 96 ATGCACCTTCCAAAGTGAAGAGAGCGCCCTGATCCCTCAGCGTTTCATGAGCTCTTGTCC 155
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
Db 156 ACGGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 215
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
Db 216 CTAATCTTCTCTTTGGGACCACTTGGCTGAGAGACCTGTGCTGAGGCTGCCAGGCC 275
QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db 276 AGCGGCATCTGCTGTGTACCACTCCCTCTTTGCTCTGCGCCAGGAGGACCTGCTCTGC 335
QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db 336 TGGTTCCCGCCGAGCCACATCTTCTCGTGAGGATGCCAGCACCCCAAGTCTCTGCTGTAC 395
QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
Db 396 AGGATTCGCTTTTACTTCCCAATGTTGGCTGAGAGAGTCCACCGCTTCCGGGCTA 455
QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
Db 456 CGCAAGGATTTGGCCAGTGTATCTTGAACCTGTCAGTCTCTGAGACACCTCTTTGCCAG 515
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160

```

516 CACCGAGTGCCTGGTGGGGGCGCTCCCGGTGGGCTCAGTCTCAAGAGCAGGCT 575
161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
576 GAGTCTCAGCCTGGCGGTGTGGACCTGGCCCGGATGGCGGAGAGCAGGCCCGAGCGG 635
181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
636 CCGGAGAGCTGCTGAAGACTGCTAGCTACAGGCTGCTACCCCAAGCCTCGCGAC 695
201 ValIleGlnGlnAnspPheValThrArgArgAlaGlyArgThrValValLeuAla 220
696 CTGATCCAGGGCCTGAGTCTCGCGGAGGGCTATTCCGAGGACGGTGGCGAGAGCC 755
221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLe 240
756 CTGCC-GGCGTGGCGGCTGCGAGGACCGGCACTCGCTCATGGCCAAAGTACATCAT 814
240 uAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260
815 GGACCTGGAGCGGCTGGATCCAGCGGGGCCCGGAGACCTTCACAGTGGGCTCCCTGG 874
260 yAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePr 277
875 GGGCCTTGGTGGCCACGACGAGGCTGGGGCTGCTCGCGGTGGTGTGACGGCGGATCGC 934
277 oTriPserSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAs 296
935 CTGGACCCAGGAGAACAGGAGGTCTCCAGGCCCTTCGCGACTTTCAGAAATCGTAGA 994
296 pValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrVa 316
995 CATTAGCATCAAGCAGGCCCCCGCGGTGGCCCGGAGGACACCGCTGGTCACTGT 1054
316 lThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSe 336
1055 TACCAGGACAGACACCAAGATTTTAGAGCGCGAGTTCCAGGGCTGCCGAGGCTCTGTC 1114
336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCy 356
1115 GTTCTGGCGCTCGTGGACGGCTACTTCGGGTGACCCAGGACTCCCGACACTTCTTCTG 1174
356 sLysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProI 376
1175 CAAGGAGGTGGCACCGCGAGGCTGCTGGAGGAGTGGCGGAGCAGTGCACGCGCCCAT 1234
376 eThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrI 396
1235 CACTCTGGACTTTGGCATCAACAAGCTCAAGACTGGGGGCTCACGCTCCTGGCTCTATGT 1294
396 eLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrPr 416
1295 TCTCGCGCGAGCCCCAGGACTTTGACAGCTTCTCCTCACTGTCTGTGTCAGAACCC 1354
416 oLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLe 436
1355 CTTTGGTCTCTATTATAGGCTGCTCATCCGGCGGAGCCCCACAGGAACTTCTTCT 1414
436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaCysTyrAsnSe 456
1415 GGTGGCTCAGCCGACCCACAGCAGTCTTCGAGAGGCTTCCTGGCAACCTGCTGGATGG 1474
456 rGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLy 476
1475 GGGGCTGACGTAGTGGGTGGAGTACCCCTCACTTCTGCTGTATCCCCAGACCCAA 1534
476 sGluLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCy 496
1535 AGAAAGTCCAACTGATGTCGTGTCAGAGAGGTACAGGCCACCCACATCATCTCTGGT 1594
496 sSerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSe 515
1595 TCAGGCCCCAATCCCAATACCAGTCTGAGTCAGATGACATTTTCAAGATCCCTGCTGACAG 1654

515 rleuGlnTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerAr 535
1655 CTTGAGTGGCATGAGAACCTGGGCGCATGGGTCTTCCACAGATTTTACCGGGGTGTGG 1714
535 gArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSe 555
1715 CCATGAGTGGTGGATGGGAGGCCCGAAGACAGAGAGGTGCTGCTCAAGGTCATGGATGC 1774
555 rArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575
1775 CAAGCACAGAAGCTCATGGAGTCACTCTGGAAGCAGCAGCTTGATGAGCAAGTGTGTC 1834
575 rTyrProHisLeuValLeuHisGlyValCysMetAlaGlyAspSerIleMetValGl 595
1835 GTACCGGCTATCTGGTGTGCTCCACGGGTGTGCTGCTGGAGACAGCACCATGTGTGA 1894
595 nGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSe 615
1895 GGAATTTGTACACCTGGGGGCCATAGACATGATCTCTGGAAAACGTGGCCACTGTGTGCC 1954
615 rAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAs 635
1955 AGCCAGCTGGAAGCTGCAGGTGCTCAAACAGCTGGCTACGCCCTCAACTATCTGGAGA 2014
635 pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGl 655
2015 CAAGGCTTGGCCCCATGCAATGTCTCTGCCGGAAGTGTCTCTGGGTCCGGAGAGGGGCGC 2074
655 yAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSe 675
2075 TGATGGAGACCGCCCTTCATCAAGCTGAGTGACCTCGGGGTACGCCCGCTGTGTAAAG 2134
675 rLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGl 695
2135 CCTGAGATGTCTACCGCAGAGATCCCTGGGTGGCCCGCGAGTGTCTCCGGAGGCGCA 2194
695 nThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGl 715
2195 GACACTTAGCTTGAAGCTGACAAAGTGGGGTCTCGGCGCCACGGTCTGGGAAGTGTTAG 2254
715 nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAs 735
2255 TGGCGTACCATGCCCCATCAGTCCCTGGATCTCTGCTAAGAAACTCCAAATTTATAGAGA 2314
735 pGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMe 755
2315 CCGGAGCAGCTGCGCGGCCCAAGTGGACAGAGCTGGCCCTGCTGATTCAACAGTGCAT 2374
755 tAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgLeuAsnGlyLe 775
2375 GGCCTATGAGCCGGTCCAGAGGCGCTCTCTCCGAGCGCTCATCTGTGACCTCAATAGCCT 2434
775 uIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 795
2435 CATCTCTCAGACTATGAGCTCTCTCAGACCCACACCTGGTGGCCCTGGCACCTCGTGA 2494
795 pGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGl 815
2495 TGGGCTGTGG---AATGGTGCCAGCTCTATGCTGCCAAGACCCACGATCTTCGAGGA 2551
815 uArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCy 835
2552 GAGACCTCAAGTACATCTCACAGCTGGGCAAGGCAACTTTTGGCAGCGTGGAGCTGTG 2611
835 sArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHi 855
2612 CCGCTATGACCCGTAGCCCAATACAGGTGGCTGTGGTGGCCGTGGAACAGCTGACAGA 2671
855 sSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHi 875
2672 CAGCGGCCACAGACAGCAGAGGGGACTTTTACGCGGAGATTCAGATCCTCAAAGCACTGCA 2731

875 sSerAspPheIleValIleTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuAr 895
 2732 CAGTGATTTCAITGTCAGTAATCGTGTGTACAGTATGCGCGGCGGACAGAGCTGCG 2791
 895 gLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgG 915
 2792 GCTGGTCAATGGAGTACCTGCCAGCGGCTGCTTGGCGGACTTCCCTCAGCGGCGACCGCG 2851
 915 y--LeuHisThrAspArgLeuLeuLeuPheAlaTyrGlnIleCysLysGlyMetGluTy 934
 2852 GCGCTCGATGCGCGGCGCTCTCTCTATTCTCGCAGATCTGCAAGGGCATGAGAT 2911
 934 rLeuGlyAlaArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlu 954
 2912 CCGTGGCTCCGCGGCTGCGTGCACCGGACCTGGCGCGCGGCGGCGGCGGCGGCGGCGG 2971
 954 rGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAs 974
 2972 CGAGGCACACGTCAGATCGCTGACTTGGCGCTAGCTAAGCTGCTGCGGCTTGACAAAGA 3031
 974 pTyrTyrValValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLe 994
 3032 CTACTACGTGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3091
 994 uSerAspAsnIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGl 1014
 3092 CTGGGACACATCTTCTCTCGCAGTCAGACGCTGAGAGCTTGGGCTGCTGCTGACGA 3151
 1014 uLeuPheThrTyrCysAspLysSerCysSerProSerProSerAlaGluPheLeuArgMetG 1034
 3152 GCTCTTACCTACTGCGACAAAGCTGACGCCCTCGGCGGAGTTCCTGGGATGATGGG 3211
 1034 yProGluArgGluGlyProProLeuCysArgLeuLeuLeuLeuAlaGluGlyArgAr 1054
 3212 ATGTGAGCGGATGTCCCGCCCTCTGCCCTCTTGGAACTGCTGGAGAGGCGGCGAG 3271
 1054 gLeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAl 1074
 3272 GCTGGCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3331
 1074 aProGluProProProProProProProProProProProProProProProProProPro 1094
 3332 CCTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3391
 1094 pArgGlyArgProGly 1099
 3392 GAGCGGAGCGGCGG 3407

RESULT 7
 AR270492
 LOCUS 3620 bp DNA linear PAT 10-APR-2003
 DEFINITION Sequence 1055 from patent US 6500938.
 ACCESSION AR270492
 VERSION AR270492.1 GI:29701726
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3620)
 AUTHORS Au-Young,J. and Sellhauer,J.J.
 TITLE Composition for the detection of signaling pathway gene expression
 JOURNAL Patent: US 6500938-A 1055 31-DEC-2002;
 FEATURES
 Location/Qualifiers
 1..3620
 /organism="unknown"

BASE COUNT 670 a 1152 c 1076 g 722 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7,27e-273 Length: 3620
 Score: 4597.00 Matches: 895
 Percent Similarity: 86.80% Conservative: 65
 Best Local Similarity: 80.92% Mismatches: 123

Query Match: 78.45% Indels: 24
 DB: 6 Gaps: 7
 US-09-397-967A-16 (11-1099) x AR270492 (11-3620)
 QY 1 MetAlaProSerSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
 DB 96 ATGGCACCTCCAAAGTGAAGAGAGCGCCCTGATCCCTCAGCGTTCATGCGACCTCTTGTTC 155
 QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
 DB 156 ACGAGGCTGTGGCTCTGATGTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 215
 QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
 DB 216 CTATCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 275
 QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
 DB 276 AGCGGATCTGCTGCTGTACCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 335
 QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
 DB 336 TGGTTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 395
 QY 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120
 DB 396 AGGATTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 455
 QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
 DB 456 CGAAGGATTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 515
 QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
 DB 516 CACCGGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 575
 QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
 DB 576 GAGTGTCTCAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 635
 QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
 DB 636 CCGGAGAGCTGCTGAAGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 695
 QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgIleArgThrValValLeuAla 220
 DB 696 CTGATCCAGGCGCTGAGCTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 755
 QY 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLe 240
 DB 756 CTGCG-CCGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 814
 QY 240 uAspLeuArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260
 DB 815 GGACCTGGAGCGGCTGGATCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 874
 QY 260 yAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePr 277
 DB 875 GGCCTTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 934
 QY 277 oTrpSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAs 296
 DB 935 CTGGACCCAGGAGAACACAGGAGTCTCTCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 994
 QY 296 pValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrVa 316
 DB 995 CATTAGCATCAAGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1054
 QY 316 lThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSe 336
 DB 1055 TACGAGGACAGACAAACAGATTTTAGAGGCGGAGTTCACAGGCGGCGGCGGCGGCGGCGGCG 1114

QY 336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCy 356
Db 1115 GTTCGTGGCGCTCGTGGACGGCTACTTCGGCTGACCCAGGACTCCAGCACTTCTTCG 1174
QY 356 slyGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIl 376
Db 1175 CAAGAGGTGGCACCGCGAGCGCTGTGAGGAAGTGGCGAGCAGTCACGGGCGCCAT 1234
QY 376 eThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIl 396
Db 1235 CACTCTGGACTTTGCCATCAACAGCTCAAGACTGGGGGCTCAGCTCTGGCTCTCTATG 1294
QY 396 eLeuArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrPr 416
Db 1295 TCTCGCGCGACCGCCCGAGGACTTTGACAGCTTCTCTCACTGTGTGTCCAGAACCC 1354
QY 416 oLeuGlyProAspTyrLysGlyCysIleuLeuArgGlnAspProSerGlyAlaPheSerIe 436
Db 1355 CTTTGGTCTGTATTATAAGGGCTGCTCATCCGGCGGCGAGCCGCACAGGAACCTTCTCTCT 1414
QY 436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaCysTrpAsnSe 456
Db 1415 GGTGGCTCTACCGACCCCGACAGCTCTTCGAGAGCTCTCGGAACCTGTGGATGG 1474
QY 456 rGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaProArgProly 476
Db 1475 GGGGCTGCACGTAGATGGGTGGCGTACCTCACTTCTCTGTATATCCCGACGCCAA 1534
QY 476 sGluLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCy 496
Db 1535 AGAAAGTCCAACTGTATGGTGTGGTGCAGAGAGGTGCACAGCCCGACATCATCTCTGTGT 1594
QY 496 sSerPro---SerCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSe 515
Db 1595 TCAGCCCCAATCCCAATACCAGCTGAGTCAGATGACATTTTCAAGATCCCTGTGTGACAG 1654
QY 515 rLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerAr 535
Db 1655 CTTGAGTGGCATGAGAACCTGGGCGCATGGGTCTTCACCAAGATTTACCGGGGGTGTG 1714
QY 535 gArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSe 555
Db 1715 CCATGAGTGTGTGATGGGAGCGCCGAAAGACAGAGGTGTGTCTGAAGGTTCATGGATGC 1774
QY 555 rArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575
Db 1775 CAAGCACAAAGAACTGCATGGAGTCATTTCTGGAAGCAGGAGCTTGATGAGCAAGTGTG 1834
QY 575 rTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGl 595
Db 1835 GTACCGGCATCTCGTGTCTCCACGGGTGTGTCATGGTGGAGACAGCACCATGTTGCA 1894
QY 595 nGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSe 615
Db 1895 GGAATTTGTACACCTGGGGCGCATAGACATATATCTGGAACACGTGGCCACCTGTGTGCC 1954
QY 615 rAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAs 635
Db 1955 AGCCAGCTGGAAAGCTGCAAGTGGTCAACAGCTGGCCTACGCCCTCAACTATCTGGAGGA 2014
QY 635 pLyGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGl 655
Db 2015 CAAAGGCTGCCCCATGCAATGTCTCTGCCCGGAAGGTGCTCTCTGGTTCGGGAGGGGC 2074
QY 655 yAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSe 675
Db 2075 TGATGGAGCGCGCGCTTTCATCAAGCTGAGTCACCTGGGGTCAGCGCCCGCTGTGTTAAG 2134
QY 675 rLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGl 695
Db 2135 CTTGAGATGCTCAACGACAGGATCCCTGGGTGGCCCCCGAGTGTCTCCGGGAGGCGCA 2194
QY 695 nThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGl 715

Db 2195 GACACTTACCTTGGAGAGCTGACAACTGGGGCTTCGGGCGCCACGGTCTGGAGAGTGTITAG 2254
QY 715 nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAs 735
Db 2255 TGGCGTCAACCTACGATCGCATCAGTGGCGCTGATCCTCTGAAGAACTCCAATTTTATGAGA 2314
QY 735 pGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysSe 755
Db 2315 CCGGACAGAGCTGGCGGCGCCCAAGTGGACAGCTGGCCCTGTGATTCACAGTGCAT 2374
QY 755 tAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775
Db 2375 GSCCTATGAGCGGTCAGAGGCCCTCTTCGAGCGCTCATTCGTGACCTCAATAGCCT 2434
QY 775 uIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 795
Db 2435 CATCTCTTCAGACTATGAGCTCTCTCAGACCCACAGCTGGTGGCCCTGGCACCTCGTA 2494
QY 795 pGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGl 815
Db 2495 TGGGCTGTGG---AATGGTGCCAGCTCTATGCCCTGCCAAGACCCACGATCTTCGAGA 2551
QY 815 uArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCy 835
Db 2552 GAGACACCTCAAGTACATCTCACAGCTGGGCAAGGGCAACTTTGGCAGCGTGGAGCTGTG 2611
QY 835 sArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHi 855
Db 2612 CGGCTATGACCGCTAGGCGCAATACAGGTGCCCTGGTGGCGCTGAAACAGCTGCAGCA 2671
QY 855 sSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHi 875
Db 2672 CAGCGGGCGAGACAGCAGAGGAGCTTTACGGGAGATTTCAGATCCTCAAGAGCTGCA 2731
QY 875 sSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuAr 895
Db 2732 CAGTGAATTTCAAGTATCGTGGTGTGAGCTATGGCCGGCGCCGAGAGCTGCG 2791
QY 895 gLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGl 915
Db 2792 GCTGTGATGAGTACCTGCCCGCGGTGTGTGGCGACTTCTGCGAGCGCACCGCGC 2851
QY 915 y---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTy 934
Db 2852 GCGCTCGATGCCCGCGCTCTTCTATTCTTCGAGATCTGCAAGGGCATCGAGTA 2911
QY 934 rLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSe 954
Db 2912 CCTGGGCTCCCGCGCTGCGTGACCGGACCTGGCCGCCCAACATCTCTCGTGAGAG 2971
QY 954 rGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAs 974
Db 2972 CGAGCACACGCTCAAGATCGCTGACTTCGGCTAGCTAAGCTGTCTGGCTTCGCGCTTGA 3031
QY 974 pTyrTyrValValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerIe 994
Db 3032 CTACTACGTGTGCGGAGCGAGCGCCAGAGCCCATTTCTGGTATGCCCCGAATCCCT 3091
QY 994 uSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGl 1014
Db 3092 CTGGACACATCTTCTCTGCCAGTCAGAGCTGAGAGCTTCGGGGTGTCTGCTGATAGA 3151
QY 1014 uLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGl 1034
Db 3152 GCTCTTCACTTACCTGCGACAAAAGCTGACGCCCTCGGCGCGAGTTCCTCGCGATGATGG 3211
QY 1034 yProGluArgGluGlyProProLeuCysArgLeuLeuLeuAlaGluGlyArgAr 1054
Db 3212 ATGTGAGCGGATGTCCCCCGCCCTCTGCGCTCTTGGAACTGTGGAGGAGGCGCAG 3271
QY 1054 gLeuProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAl 1074

815 GGACCTGGAGCGGCTGGATCCACCGCGGGCCGCGAGACCTTCCACGTGGGCTCCCTGG 874
260 yAla-----GlnGluGluProGlyLeuArgValAlaGlyAspAsnGlyLePr 277
875 GGCCTTGGTGGCCAGCAGCGGCTGGCTCGCGGTGGTGGTGGAGCGGCGATCGC 934
277 ofrpSerSerAnApp--GluLeuPheGlnThrPheCysAspPheProGluLeValAs 296
935 CTGGACCCAGGAGAAACAGAGGTCCTCCAGCCCTTCGGACTTTCAGAAATCGTAGA 994
296 pValSerIleAnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValrVa 316
995 CATTTAGCATCAACAGAGGCCCGCGCTTGGCCGCGCGAGAGCACCGGCTGGTCACTGT 1054
316 lThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSe 336
1055 TACCAGGACAGAAACAGATTTTAGAGGCCGAGTTCCAGGGCTGCCGAGGCTCTGTC 1114
336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCy 356
1115 GTTCGTGGCGCTCGTGGAGCGGTACTTCGGGTGACCACGGACTCCAGCACTTCTCTG 1174
356 sIysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIl 376
1175 CAAGGAGGTGGCACCGCGGAGGTGCTGGAGGAAAGTGGCGGAGCAGTGCACGGGCCCAT 1234
376 eThrLeuAspPheAlaIleHisIleLeuLysAlaAlaGlySerLeuProGlyThrTyrIl 396
1235 CACTCTGGACTTGGCATCAACAGCTCAAGACTGGGGGCTCAGGTCTCGGTCTCTATGT 1294
396 eLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrPr 416
1295 TCTCGCGCGACGCCCGGAGGACTTTGACAGCTTCTCTCTCACTGTGTGTCCAGAACCC 1354
416 oLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerIe 436
1355 CTTGGTCTCTGATTATAAGGGTGTGCTATCGGGGCGAGCCCGCAGAGAACCTTCTTCT 1414
436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaCysTyrAsnSe 456
1415 GGTGGCTCAGCCAGACCCCAAGAGCTTTCGAGAGCTCTCGGCAACCTGCTGGATGG 1474
456 rGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProly 476
1475 GGGGTGCACGTAGATGGGTGGAGTGACCTCACTTCTCTCTGTATCCCCAGACCCAA 1534
476 sGluLysSerAnLeuIleValValArgGlyCysAnProAlaProAlaProGlyCy 496
1535 AGAAAGTCCAACTGTATGATGCTGCTGTCAGAGAGGTACAGCCCGCAGCATCATCTTGT 1594
496 sSerPro--SerCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSe 515
1595 TCAGGCCCCAATCCCAATACCAGCTGAGTCAGATGACATTTTCAAGATCCCTGCTGACAG 1654
515 rLeuGluTrpHisGluAnLeuGlyHisGlySerPheThrIysIlePheArgGlySerAr 535
1655 CTGGAGTGGCATGAGAACCTTGGGCATGGTCTTCCACAGATTATCCGGGGCTGTGC 1714
535 gArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSe 555
1715 CCATGAGTGGTGGATGGGGAGCGCCGAAAGACAGAGGTGCTGTCTGAAGTTCATGGATGC 1774
555 rArgHisArgAnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575
1775 CAAGCACAGAAACTGCATGGAGTCAATTCCTGGAAGCAGCGAGCTTGTGAGCAAGTGTTC 1834
575 rTyrProHisLeuValLeuHisGlyValCysMetAlaGlyAspSerIleMetValGl 595
1835 GTACCGGCATCTCTGCTGCTCCAGCGGTGTGCATGCTGGAGACAGCACCATCGTGCA 1894
595 nGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSe 615
1895 GGAATTTGTACACCTGGGGCCATAGACATATCTGCGMAAACGTGGCCACCTGGTGGC 1954

QY 615 rAlaSerTrpLysLeuGlnValThrIysGlnLeuAlaTyrAlaLeuAsnTyrIleuGluAs 635
Db 1955 AGCCAGCTGGAAAGCTGCAGGTGGTCAAAACAGCTGGCTTACGCCCTCAACTATCTGGAGA 2014
QY 635 pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGl 655
Db 2015 CAAAGGCTGCCCATGGCAATGTCTCTGCCCGGAAGGTGCTCTGGCTCGGAGGGGGC 2074
QY 655 yAspGlyAnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSe 675
Db 2075 TGATGGAGCGCCGCTTTCATCAAGCTGAGTCACTGGGGTTCAGCCCGCTGTGTAAAG 2134
QY 675 rLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGl 695
Db 2135 CTGGAGATGCTCAACGACAGGATCCCTGGGTGGCCCCCGAGTGTCTCCGGAGGGCGCA 2194
QY 695 nThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGl 715
Db 2195 GACACTTAGCTTGGAAAGCTGACAGTGGGGCTTCGGCGCCACGGTCTGGGAAGTGTTAG 2254
QY 715 nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAs 735
Db 2255 TGGCGTCAACATGCCATCAGTGCCTGGATCCTCTGAAGAACTCCAATTTATGAGGA 2314
QY 735 pGlnGlyGlnLeuProAlaLeuLysTrpThrGlnLeuAlaGlyLeuIleThrGlnCysMe 755
Db 2315 CCGGACAGCAGTGC CGGCCCCCAAGTGACAGAGTGGCCCTGCTGATTCACACAGTGCAT 2374
QY 755 tAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775
Db 2375 GGCCTATGAGCGGTTCAGAGGCCCTCTCTCGAGCGCTCATTCGTGACCTCAATAGCT 2434
QY 775 uIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 795
Db 2435 CATCTCTTCAAGTATGAGTCTCTCTCAGACCCACACCTGGTGGCTGGCACCTCGTGA 2494
QY 795 pGluLeuCysValAlaAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGl 815
Db 2495 TGGGTGTGG---AATGGTGGCCAGCTCTATGCTGCCAAGACCCCAAGTCTTCGAGGA 2551
QY 815 uArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCy 835
Db 2552 GAGACACCTCAAGTACATCTCACAGTGGGCAAGGGCAACTTTTGGCAGCGTGGAGCTGTG 2611
QY 835 sArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHi 855
Db 2612 CCGCTATGACCGCTAGCGGCAATACAGGTGCCCTGGTGGCGGTGAAACAGCTGCAGCA 2671
QY 855 sSerValProAspGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHi 875
Db 2672 CAGCGGGCAGACCCAGCAGAGGGACTTTTCAGCGGAGATTCAGATCCTCAAGACACTGCA 2731
QY 875 sSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuAr 895
Db 2732 CAGTGATTTTCATTGTCAAGTATCGTGGTGTAGCTATGGCCCGGCGCCAGAGCTGCG 2791
QY 895 gLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGl 915
Db 2792 GCTGGTTCATGAGTACCTGCCCGGCTGTCTGGCGACTTCCTGTCAGCGGCACCGCGC 2851
QY 915 y---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTy 934
Db 2852 GCGCTCGATGCCAGCGCCCTCTCTCTATTCGCGAGATCTGCAAGGGCATGGAGTA 2911
QY 934 rLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSe 954
Db 2912 CTGGGCTCCCGCGCTGCTGCGACCGGACCTGGCGCCCGCAAAACATCTCTCTGGAGAG 2971
QY 954 rGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAs 974
Db 2972 CGAGGCACACGTCAAGATCGTACTTCGGCTAGCTAAGCTGCTGCCGCTTGACAAAGA 3031

```
QY 974 pYrVrValValArgGluProGlyGlnSerProIlePheThrTyrAlaProGluSerLe 994
Db 3032 CTACTAGTGGTCCGAGCCAGCCAGCCCAATTTCTGATGCCCCCAATCCCT 3091
QY 994 uSerAspAenIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGl 1014
Db 3092 CTCGACACATCTTCTCTCCAGTACAGAGTCTGGAGTCTGGGGTCTCTGTACGA 3151
QY 1014 uLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGl 1034
Db 3152 GCTCTTACACTACTGACAAAAGCTGACGCCCTCGCCGAGTCTCGGGAGTGGG 3211
QY 1034 yProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgAr 1054
Db 3212 ATGTAGCGGAGTCTCCCGCCCTCTGCCGCTCTTGAAGTCTGTGAGAGGCCAGAG 3271
QY 1054 gLeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAl 1074
Db 3272 GCTGCCGCGCCCTCTGCTGCCCTGCTGAGGTG----- 3305
QY 1074 aProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTr 1094
Db 3306 -----AGCCCGAGGCTAGCTCAGTTTCCAGTCTGTAGAT-----TG 3346
QY 1094 pArgGlyArgProGly 1099
Db 3347 GCGCGGGTCTCGGC 3362

RESULT 9
AC073700
LOCUS
DEFINITION Mus musculus clone RP23-162012, WORKING DRAFT SEQUENCE, 16
unordered pieces.
ACCESSION AC073700
VERSION AC073700.1 GI:8810317
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 218208)
DOE Joint Genome Institute.
Sequencing of Mouse
Unpublished
2 (bases 1 to 218208)
DOE Joint Genome Institute.
Direct Submission
Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 1799428
Center clone name: RPCI-23_162012
-----
Summary Statistics
Consensus quality: 204328 bases at least Q40
Consensus quality: 210800 bases at least Q30
Consensus quality: 212488 bases at least Q20
Estimated insert size: 222000; agarose-fp estimation
Estimated insert size: 216708; sum-of-contigs estimation
Quality coverage: 9.37 in Q20 bases; agarose-fp estimation
Quality coverage: 9.6 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
```

```
* be preserved.
* 1 1092: contig of 1092 bp in length
* 1093: gap of unknown length
* 1193: contig of 1373 bp in length
* 2565: gap of unknown length
* 2656: gap of unknown length
* 2866: contig of 1046 bp in length
* 3712: gap of unknown length
* 3812: contig of 1744 bp in length
* 5555: gap of unknown length
* 5556: gap of 3359 bp in length
* 9014: gap of unknown length
* 9115: contig of 3410 bp in length
* 12525: gap of unknown length
* 12525: contig of 5691 bp in length
* 18315: gap of unknown length
* 18316: contig of 6884 bp in length
* 25299: gap of unknown length
* 25300: gap of unknown length
* 25400: contig of 9459 bp in length
* 34858: gap of unknown length
* 34959: contig of 13153 bp in length
* 48111: gap of unknown length
* 48112: contig of 11464 bp in length
* 59675: gap of unknown length
* 59676: contig of 19464 bp in length
* 59775: gap of unknown length
* 59776: contig of 23420 bp in length
* 79240: gap of unknown length
* 102759: contig of 27114 bp in length
* 102760: gap of unknown length
* 102859: contig of 27114 bp in length
* 12973: gap of unknown length
* 12974: contig of 34115 bp in length
* 130074: gap of unknown length
* 164189: contig of 53920 bp in length.
* 164289: contig of 53920 bp in length.

FEATURES
Location/Qualifiers
source
1..218208
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone_lib="RPCI mouse BAC library 23"
BASE COUNT 51066 a 56856 c 56653 g 52132 t 1501 others
ORIGIN

Alignment Scores:
Pred. No.: 1.2e-239 Length: 218208
Score: 4091.50 Matches: 1072
Percent Similarity: 32.75% Conservative: 2
Best Local Similarity: 32.69% Mismatches: 24
Query Match: 69.82% Indels: 2188
DB: 2 Gaps: 24

US-09-397-967A-16 (1-1099) x AC073700 (1-218208)
QY 1 MetaAlaProSerGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
Db 110153 ATGCACCTCCAGTGAGGACACCTCTGATCCCTCAGGCTCTTCAGCTCTCATCC 110212
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
Db 110213 TCAGAGCAGGAGCCCTGCATGTCTCTTCTCCCGGGGACCTGGGCTCCCGAGCGA 110272
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
Db 110273 TTGTCAATCTCTTTGGGGGACTACTTGGCTGAGGATTTATGTGTCGAGCTGCCAAGGCC 110332
QY 61 CysGlyLe----- 63
Db 110333 TGTGTGA-GTGGGTCCCTAGATGACCATGCTATGTAACATCTATTGAAAGG 110391
QY 63 ----- 63
Db 110392 GGATGCTGGGGTTTTGAGGGGTGTGGAATTTTGTAAAGGAGGATCATGGGATCAGCCCG 110451
QY 64 -----LeuProValTyrHis 68
```

Db	110452	TGTA	TTT	ATCTTCAGGGTAGTATTTCTATTCAGTCTATAGTAGAGGCTCTTGAACCCAGA	110691
QY	69	SerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPheProProSerHisIlePhe	88		
Db	110512	TGCCTTTTCGCTCTGGCCACTGAGGACTTCTTGTCTGTCTTCCCCAGCCACATCTTC	110571		
QY	89	CysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg	103		
Db	110572	TGCATAGAGGACGTGGACACTCAAGTCTTTGGTCTACAGGCTACGGTAGGAAGTGAACCC	110631		
QY	103	-----	103		
Db	110632	TTGATTTATCTCAGGGTAGTATTTCTATTCAGTCTATAGTAGAGGCTCTTGAACCCAGA	110691		
QY	103	-----	103		
Db	110692	GAGGTTGGGCTGTTTGTCTAGCCACACTGCACCTCAGTGCACCCAACTTGGGCTCTGGG	110751		
QY	104	-----	104		
Db	110752	AAGCCAGAGCTGAGTGGGCCACCCCATCTCTGAAACCCACAGCTTTTATTTCCCTGACTGG	110811		
QY	110	PheGlyLeuGluThrCysHisArgPheGlyLeuArgLysAspLeuThrSerAlaIleLeu	129		
Db	110812	TTTGGGCTGGAGACATGTCACCGCTTTGGCTGGCGCAAGATTTGACCAAGTGCCTT	110871		
QY	130	AspLeuHisValLeuGluHisLeuPheAla	139		
Db	110872	GACTTACACGTTTTAGAACATCTCTTTGCTCAGGTTGGTATTTCTTCGGTGGGAGGCTC	110931		
QY	139	-----	139		
Db	110932	AACCCAGGATGGGGACCCCAAGGTAAACAATCCAGCACCCAGGACGGAACTCACCCT	110991		
QY	140	-----	140		
Db	110992	AACCAACCCTTCTGTGGCTGACCCAGCACCGCAGTCACTGTGGTGGGCGGCTCCCGGT	111051		
QY	152	IgLYLeuSerMetIysGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaG	172		
Db	111052	GGGCTTAGCATGAAGGAGCAGGAGAGTTCCTGAGCTGGCGTGTGGACTTGGCCCA	111111		
QY	172	nMetAlaArgGluGlnAlaGlnArgProGlyIleLeuLeuIysThrVal	188		
Db	111112	GATGGCTCGTGAGCAGGCCACGCGCCAGGAGAGTCTGCTGAAGACCGTCAGGTGAGGGCC	111171		
QY	188	-----	188		
Db	111172	TCGGGCCACCCCTTCGTGGTCTTCTTTGGCTGGGATCTAGGGCCCTATCGTTTTTCCCTT	111231		
QY	188	-----	188		
Db	111232	ACTGGAACACATACCTCTTCCCAACTGTGCAGCCACCCTCTCGTCAGCGTGGCCCTG	111291		
QY	188	-----	188		
Db	111292	CCGAAAGTCCCTATGTCTGTCTGTGTCTTCGTGGGCCCCCTCACCACTTAAGAGC	111351		
QY	189	-----	189		
Db	111352	TCGGGTGCCCCCCCCCAGTGTACAAAGCTGTCTCCGCCAGCCCTGGCGGATGTGA	111411		
QY	202	IeGlnGlyGlnAsnPheValThrArgArgGilleArgArgThrValValLeuAlaLeuL	222		
Db	111412	TCCAGGGCCAGAACTTCGTGTGACACGACGGCATCCGACAGACCGTGTCTTGGCGCTGC	111471		
QY	222	euproCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLeuAsp	241		
Db	111472	G-CGGTGTGGTGGCTGCGAGCCGACCGCTACGCGCTCATGGCCAGTATATTTGGAC	111530		
QY	242	LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAla	261		

Db	111531	CTGGAGCGCTACATCCACGGCCACCAACCGAGACCTTCGCTGTGGGGCTCCCGGGCGGC	111590
QY	262	GingluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrpSerSerAsn	281
Db	111591	CAGGAGGACCGGGGCTTCGTGTGGCGGGGACACGGCATCTCTGGAGCTCCGGG	111650
QY	282	AspGlu-----	283
Db	111651	GACCA-GGAGGTGTGGGCTGGGCTTGAGAAGCTGGGTCCGCGGGAGGCGTGGCGGT	111709
QY	283	-----	283
Db	111710	TAGTAAAGAAGCCCGTCCGAACAATCAGAAAAGTGCTAGTCCGGCGGGACTGATTGCA	111769
QY	283	-----	283
Db	111770	CTACGGGCAGGGGGCGGGTCAATTAAAGTGGTCCGGGATGGAGCTACCCAGGGA	111829
QY	283	-----	283
Db	111830	CTGGACCTGGGGCGTGACTAGAAAGGTGAGACTAGATAGGGGCGGGACGACAGAA	111889
QY	283	-----	283
Db	111890	CGCATAGTATGGATTGGATCGGTGGGTGTGACCCGACCAGGCCCTGTCTCAGTCCCGCTAAC	111949
QY	283	-----	283
Db	111950	CTCGTCCCGGACCCCATCCACCCCTCCCGTAAAGTGGAGGTGGGAAGGAGGAAG	112009
QY	283	-----	283
Db	112010	AAGGGAGGCTGTCTGGTCAAGTCTGTAACTCTCGCTTCGCTCTCTCTTTCT	112069
QY	283	-----	283
Db	112070	CTCTTTTTTAAATGGTTTTGCGACACAGGGTTTCTCTGTATAGCTCTGGCTGCTCTGGA	112129
QY	283	-----	283
Db	112130	ACTCATTGTAGACCAGCTGGCGCTGGAACTCAGAAATCGCGCTCTGCTGCTCTCGA	112189
QY	283	-----	283
Db	112190	GTGCTGTATTAAAGGCGTGCACCTACACCCAGCTACTGCCTGTCTCTTCTCTCTCAG	112249
QY	284	LeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaPro	303
Db	112250	CTTTTCAGACCTTCTGTGACTTTCCGAAAATCGTGGATGTGAGCATCAAGCAGGCCCA	112309
QY	304	ArgValGlyProAlaGlyGluHisArgIleuValThrValThrArgMetAspGlyHisIle	323
Db	112310	CGTGTGGTCCGGCAGGGGAACACCGGTGTGTCCCGTCACAGGATGGACGGCCACATC	112369
QY	324	Leu-----	324
Db	112370	CTGGTGGTCTTGGCACCGGTATTCTGTGCTGATCTCTGGAAAGGGTGGGTTCGGT	112429
QY	325	-----GluAlaGluPheProGlyLeuProGluAlaLeuSe	336
Db	112430	GTTGCTGAAGAGTTCTCTCGCAGGAAGCGAGTTTCGGGGGTGCTCTGAGGCGCTGC	112489
QY	336	rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCys	356
Db	112490	CTTCGTGGCCCTTGTGATGGTACTTCCCGCTGATCTGCGACTCCAGGCATTATTTCTG	112549
QY	356	sLysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIle	376
Db	112550	CAAGAGGTGGCGCGCGGCTGTCTGGAGGAGGAGCGGAGCTGTGCCATTGACCCAT	112609
QY	376	eThr-----	377
Db	112610	CACGTAGGAGCTAGTTGGGCTCACACTGTGGGATGGGGCTTGGGCAGCGAGCATCTCTCG	112669

QY	378	-----LeuAspPheAlaIleHisLysLeuLysAlaA 388
Db	112670	GTTGGACCTGGCTGCTGTCTTGGCCCTAGGTTAGACTTGCCATCCACAAGCTGAAGCGC 112729
QY	388	laGlySerLeuProGlyThrTyriLeuArgSerProGlnAspTyrAspSerPheL 408
Db	112730	CTGGCTCCCTCCCAAGCACCTACATTCTCGCGCAGCCCGAGGACTATGACAGCTTTC 112789
QY	408	euleuthrAlaCysVal-
Db	112790	TTCTTACCGCCTGCGT--CCAGGTTGGTTTACTCTGGTGCCAGATGGCTTCTCTGGAGGA 112848
QY	413	-----413
Db	112849	AGTGACACCTGAGCTGGGACAGAAGAGGCGTTTCTATTTTTTCTTGAAGAGCATGGCC 112908
QY	413	-----413
Db	112909	TGTAATCTTAGAAATTAAGAGACTGAGGCGAGGATCAAGTTTACAGGCCCTGCCTGAGCTAT 112968
QY	413	-----413
Db	112969	ATAGAGAAATPATTTTTTCTATGGAGGAAAAGTTTTTTTCCACCAAGCTGAGGCTGT 113028
QY	413	-----413
Db	113029	CTTTTCATCCATGGCATTCACAGCATGAGAGGACAGAACTGACACCCATAAATCGCGTCTG 113088
QY	413	-----413
Db	113089	ACCTCCACAGTGCAAGTGACTTGCATCTACACAGAAACACACATNAAGTGAATGTAATTT 113148
QY	413	-----413
Db	113149	TAAGAACTCGAAGAGCTCAGAAACAGTCTCAGCATGGCGGGGGGGGGGAGAG 113208
QY	413	-----413
Db	113209	CATCCTTAGAGTTGGAGACAGAACTGAGCATACAGTCTTCAGAAAGACGAGAGGGT 113268
QY	414	-----GlnThrProLeuGlyP 419
Db	113269	GGGTAGCTCTGCTCCACATCTGGCTCAGGTGAGTCTCTTTTGCAGACTCTCTCTGGCC 113328
QY	419	roAspTyrLysGlyCysLeuileArgGlnAspProSerGlyAlaPheSerLeuValGlyL 439
Db	113329	CCGACTACAAGGGCTGCCCTCATCCGCGAGGACCCAGCGGGGTTTCTCCCTGGTTGGCC 113388
QY	439	eusSerGlnProHisArgSerLeuArgGluLeuLeuAlaCysTyrAsnSerGlyLeuA 459
Db	113389	TCAGCCACCCCACAGAAAGCTCGGGAGAGTGTGTGAGCTGCTGCGAATTCCTGGGCTGC 113448
QY	459	rgValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaProArgProLys-
Db	113449	GAGTAGACGGTCTGCCCTGAACCTAACATCTGCTGCGCTCCAGACCCCAAGGGTGAGT 113508
QY	476	-----476
Db	113509	CAGTTCTCTCCCTAACACCGCGCAGCTCTTGGGTGTGTAGAGAAACCAAGGACCATGG 113568
QY	476	-----476
Db	113569	CGATTAGGACAGAGAACGACTTTTGGGCGACAGTCTTTCAGACATCAAGTCAGTTATTTT 113628
QY	476	-----476
Db	113629	CAGTTGGCGCGGTTTCTTCCATGGGTGAATGCCACTAAGCGGAACCCCTGCTGGGACA 113688
QY	476	-----476
Db	113689	CATCTCCCCCTCTGTGTGTCTTGGCTAACAAAGACCAAGCGCGGCTTGGAGGAAACCCAGT 113748

QY	801	lyAlaGlnIleuTyAlaCysGlnAspProAlaIlePheGluGluArGHisLeuLysTyrl	821	QY	988	-----	988
Db	117044	GCSCCAGCTCTATGCTGCGAGACCCGCCATATTCGAGGAGAGACACCTTAAGTACA	117103	Db	118124	CCTCTCAACTAGTCAGGACTACCCACCACCACCCACAGAGTGATGCTCCACTCGGTTTA	118183
QY	821	leSerLeuLeuGlyLys-----	826	QY	988	-----	988
Db	117104	TCCTTTGCTGGCAGAGGTGAGTGGGCGGCGCATCTGGGGAGGNAACGTGGTGGTGAT	117163	Db	118184	GCCACGCCCCCATTTGTTTGGGTCCATCTCTTGACCAGTGTGAAAGCCCGTCAAG	118243
QY	827	-----GlyAsnPheGlySerValGluLeu	834	QY	988	-----	988
Db	117164	GGGTCAAGTGGACACTGCCCTCTCATCTCCACAGGCGCACTTTGGCAGCGTGGAGCTG	117223	Db	118244	GTCTCTTTTCTAGTCGCTCCCTCCGAGGGCCCTGCTTTCACGCTCTATGGGTCCGC	118303
QY	835	CysArgTyAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGln	854	QY	988	-----	988
Db	117224	TGCGCTATACCCCTTGGGGGCAATACCGGACCCCTGTGGCAGTGAACAGCTACAG	117283	Db	118304	CCCTCTTAGCGTCTTCTACTTCTCCGTCATCTTACTCTGTTCTCCGTCGCTGCTCT	118363
QY	855	HisSerValProAspGlnGlnArgAspPheGlnArgGlnIleLeuLysAlaLeu	874	QY	988	-----	988
Db	117284	CACAGCGGGCCAGACAGAGAGGACTTCCAGCGGAGATTTCAGATCCTTAAGCTCTG	117343	Db	118364	CCCTTATGACATTTAGGCCTAICTGCCAGGCACCTTTTCAGCTGCTGTCTCATCAATG	118423
QY	875	HisSerAspPheIleValLysTyArgGlyValSerTyArgPro-----	889	QY	988	-----	988
Db	117344	CACAGCGATTCATCGTCAAGTACCGGGAGTCACTATGGGCCAGGTGAGCGGCAGAG	117403	Db	118424	CTGACCCACTCAAGCCCGCTCTTAACACTTAGAGTCCCTCTCTCAGCCTGGCCACGCC	118483
QY	889	-----	889	QY	988	-----	988
Db	117404	CATCTCGGGAACGGGTCGAGTCGGCTTCTACCTTTCCAGTAGGGCCCTGTGAACA	117463	Db	118484	CACAGGTCAACAGACAGCCTCTGAGTCGCTTCTGGGGCGGTGCTAGTAGGAAGTCTCC	118543
QY	889	-----	889	QY	988	-----	988
Db	117464	AGGTCTTAATCCCATCAATCCAGCTTCTATAGTCGGAGTTGAACGAATCAATACAC	117523	Db	118544	TGTTACGCCCTTTCCTCACCCAGACTGGCCATCTTACATGACGACTCCCTCCCTCC	118603
QY	890	-----GlyArgGlnSerLeuArgLeu	896	QY	988	-----	988
Db	117524	CCAGTAGAAGGTGCTCGAGAGCTACTGTGCTGTTCGAGGTCCGACAGCTCGGTG	117583	Db	118604	AGGCCACTGTTTACACCCCTGCCACTCTCTCTACTGGCCACACACACCTCTGTCCGCG	118663
QY	897	ValMetGluTyLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGly---	915	QY	989	-----TyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTr	1006
Db	117584	GTGATGAGTACCTGCCCGCGCTCTCGAGACTCTTCGAGCGCCATCGCGCGCGC	117643	Db	118664	TCTCAGGTATGCCCGGAGTCCCTATCTGACACATCTCTTCCCGCCAATCTGACGTG	118723
QY	916	LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLys-----	930	QY	1006	pSerPheGlyValValLeuTyArgLeuPheThrTyCysAspLysSerCysSerProSe	1026
Db	117644	CTGCACACCGACCGCTACTGCTGTCTGTTGGCAGATCTCAAGGTGCGAGATCGCGG	117703	Db	118724	GAGCTTCGAGTGTGTGTACGAGCTCTTCCCTACTGACACAGAGCTGACGCCCATC	118783
QY	930	-----	930	QY	1026	rAla-----	1027
Db	117704	AGATCGGATGGGTCTCTGGGGTCCGGGATCCCGGCCACGTGCTATCATGCTCTCT	117763	Db	118784	CGC-TGTGCGTGGCTGCCCATCCCTCGGTCTCTCTTGTATCTCCAAATCCCTCTCT	118842
QY	931	-----GlyMetGluTyLeuGlyAlaArgCysValHisArgAspLeuAlaAlaArgAs	949	QY	1027	-----	1027
Db	117764	GCAGGCATGAGTACCTGGGTGCGCGCTGCTGCTACCGTGNACCTGGCTGGCGCAA	117823	Db	118843	GACCTTAGCCCTATCTCTGACCCAGCCCTTCTCTGACCTCCAGATCATCTCTCTGAC	118902
QY	949	nileLeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLe	969	QY	1027	-----	1027
Db	117824	CATCTTGTGGAGACGAGGCTCATGTGAAGATCGGACTTCGGCTCTGCTAAGCTGCT	117883	Db	118903	CTCAGTCTCTCTCGGAATCCAGCCCTTCTCTCTCATCTCCAGATCCCTCTCTGACCC	118962
QY	969	uProLeuGlyLysAspTyTyValValArgGluProGlyGlnSerProIlePheTrp--	988	QY	1027	-----	1027
Db	117884	GCCCTTGGGAAAGGACTACTAGTGTTCGGAGCCTGGCAAGCCCATCTTTTGTGA	117943	Db	118963	CAGCCCTTCCCTGACCCCGAGCCCTTCTCATGAGTCAAGCCCTTCTCTGACCTAAC	119022
QY	988	-----	988	QY	1028	-----GluPheLeuArgMetMetMetGlyProGluAr	1037
Db	117944	CAAGACCAAGACTCCAGAGGTCAAGCCCTGAGCCCACTCCCGCATCTACAGCCCTT	118003	Db	119023	AGTCTGCCCTAACTTGTCTCTCTCCACAGAGGTTCCTGAGCATGTATGGGGCTGAGCG	119082
QY	988	-----	988	QY	1037	gGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgArgLeuProPr	1057
Db	118004	AAACCGGTCCAGACCCAGCTCCACCTCAGTAGCTTTTGCCCTCGCAAGAGCCTTGCCACA	118063	Db	119083	TGAAGGACCCCGCTCTCCGCTCTCTGAGCTGCTGCGAGGGCCGCGCTCCACCC	119142
QY	988	-----	988	QY	1057	oProProThrCysProThrGlu-----	1064
Db	118064	GCAGGTGAGCCCTATAGCCAGATCCAGAACGACATGCTCAAACTCTACAGATGTACC	118123	Db	119143	ACCTCCACCTGCCCAACCGAGGTGAGGAGGACTCAAGTTTCTCAGTTTCAATTCT	119202
QY	1064	-----	1064	QY	1064	-----	1064


```
Db      119203  TGATTGACAGCAGCGGTGCTGACTGGGAATACTTTTACAGGATGAGGAGCTCAGCTTC 119262
QY      1064  ----- 1064
Db      119263  AAGGGCTTCAGAAATGGTTTTCACTTTTGTGAGCAAGGTCTTGCTATGTAGCCTTG 119322
QY      1064  ----- 1064
Db      119323  TCTGGCTGGAACCTGACATCCTCTGCTCAGTTTCTCTGACTGAGAGCAAAATGT 119382
QY      1064  ----- 1064
Db      119383  ATGTGCCACCAACAACCACTCCATATTTCAATTTTCTTTTATTATTACTTAA 119442
QY      1064  ----- 1064
Db      119443  TTTATTTTACAGTCTGAGATTTGAAGCCAGGGTCTTACTCATGCTAGGCAAGCAGC 119502
QY      1064  ----- 1064
Db      119503  ACCCTGTGTCATGATAATTCGGGGGATGGGGTTGGGGAGGAGCGGAGCGGCTTCCA 119562
QY      1064  ----- 1064
Db      119563  GGAGTGAATGGCGCCTATTGGAGAGCCAGGAGAGGTCAATCCCACTAGGAGACATGGCT 119622
QY      1064  ----- 1064
Db      119623  GTGGCTGGGTGGAGTAAGAGACAGATGTCGCGGAGATGGGGGATGTGTGCTGTC 119682
QY      1064  ----- 1064
Db      119683  ACCCTAGCCATCTGAGGCCAGGCGAGGAGTGTGCCACAGCAACTTAGTTCTGTTC 119742
QY      1064  ----- 1064
Db      119743  AGAGATCCATAAGGAGCCATGGAGAGCTGTGGAGAGAGTAGGGGCCATTAACTGACT 119802
QY      1065  -----ValGlnG1 1067
Db      119803  GGAACACCGCCTCTCAGCTCAGCCACCCACCTTGCTTCCCAACCAAGCTTCAGGA 119862
QY      1067  uLeuMetGlnLeuCysTrpAlaProdluProHisAspArgProAlaPheAlaThrIeuSe 1087
Db      119863  GCTCATGAGCTGTGCTGGGCGCCAGCCGACGACCGCCAGCTTCGGACCCCTGAG 119922
QY      1087  rProGlnLeuAspProLeuTrpArgGlyArgProGly 1099
Db      119923  CCCCCAGCTGGAGCGGCTGTGGCGTGAAGACCCGGA 119959

RESULT 10
AC073750
LOCUS   223734 bp DNA linear HTG 29-JUN-2000
DEFINITION Mus musculus clone RP23-330D8, WORKING DRAFT SEQUENCE, 28 unordered
pieces.
ACCESSION AC073750
VERSION   AC073750.1 GI:8810367
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 223734)
AUTHORS   DOE Joint Genome Institute.
TITLE     Sequencing of Mouse
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 223734)
AUTHORS   DOE Joint Genome Institute.
JOURNAL   Direct Submission
SUBMITTED (29-JUN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
```

```
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 1863672
Center clone name: RPCI-23_330D8
-----
Summary Statistics
Consensus quality: 196566 bases at least Q40
Consensus quality: 208241 bases at least Q30
Consensus quality: 210965 bases at least Q20
Estimated insert size: 208000; agarose-fp estimation
Estimated insert size: 221034; sum-of-contigs estimation
Quality coverage: 8.01 in Q20 bases; agarose-fp estimation
Quality coverage: 7.54 in Q20 bases; sum-of-contigs estimation
* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
```

```
* 1 1027: contig of 1027 bp in length
* 1028 1127: gap of unknown length
* 1128 1261: contig of 1034 bp in length
* 1262 2261: gap of unknown length
* 2262 3709: contig of 1448 bp in length
* 3710 3809: gap of unknown length
* 3810 4864: contig of 1055 bp in length
* 4865 4964: gap of unknown length
* 4965 5985: contig of 1021 bp in length
* 5986 6085: gap of unknown length
* 6086 7691: contig of 1606 bp in length
* 7692 7791: gap of unknown length
* 7792 10317: contig of 2526 bp in length
* 10318 10417: gap of unknown length
* 10418 11756: contig of 1339 bp in length
* 11757 11856: gap of unknown length
* 11857 13223: contig of 1367 bp in length
* 13224 13324: gap of unknown length
* 13324 16114: contig of 2791 bp in length
* 16115 16214: gap of unknown length
* 16215 20778: contig of 4564 bp in length
* 20779 20878: gap of unknown length
* 20879 23553: contig of 2675 bp in length
* 23554 23654: gap of unknown length
* 23654 26667: contig of 3014 bp in length
* 26668 26767: gap of unknown length
* 26768 30222: contig of 3455 bp in length
* 30223 30322: gap of unknown length
* 30323 34999: contig of 4577 bp in length
* 34900 34909: gap of unknown length
* 35000 39609: contig of 4610 bp in length
* 39610 39709: gap of unknown length
* 39710 46418: contig of 6709 bp in length
* 46419 46518: gap of unknown length
* 46519 50029: contig of 3511 bp in length
* 50030 50129: gap of unknown length
* 50130 56982: contig of 6853 bp in length
* 56983 57082: gap of unknown length
* 57083 66435: contig of 9353 bp in length
* 66436 73174: contig of unknown length
* 73175 73274: gap of unknown length
* 73275 87407: contig of 14133 bp in length
* 87408 87507: gap of unknown length
* 100239 100338: contig of 12731 bp in length
* 100339 109196: contig of 18858 bp in length
* 119197 119296: gap of unknown length
* 119297 140119: contig of 20823 bp in length
```

* 140120 140219: gap of unknown length
 * 140220 169227: contig of 29008 bp in length
 * 169228 169327: gap of unknown length
 * 169328 198979: contig of 29652 bp in length
 * 198980 199079: gap of unknown length
 * 199080 223734: contig of 24655 bp in length.

FEATURES

source
 1. 223734
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP23-330D8"
 /clone_lib="RPCI mouse BAC library 23"
 BASE COUNT 57507 a 53185 c 54134 g 56203 t 2705 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.24e-239 Length: 223734
 Score: 4091.50 Matches: 1072
 Percent Similarity: 32.75% Conservative: 2
 Best Local Similarity: 32.69% Mismatches: 24
 Query Match: 69.82% Indels: 2188
 DB: 2 Gaps: 24

US-09-397-967A-16 (1-1099) x AC073750 (1-223734)

QY 1 MetAlaProSerGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
 Db 124160 ATGCACCTCCAGTGAAGGACACCTCTGATCCCTCAGCGCTCTTGACGCTCTCATCC 124219
 QY 21 SerGluAlaGlyAlaLeuHisValLeuProProArgGlyProGlyProGlnArg 40
 Db 124220 TCAGAGCAGGAGCCCTGCATGTCTCTTCTCCCGGGGACCTGGGCTCCACAGGA 124279
 QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
 Db 124280 TTGTCAATCTCTTTTGGGGACTACTTGGCTGAGGATTTATGTGTGCGAGCTGCCAAGGCC 124339
 QY 61 CysGlyIle----- 63
 Db 124340 TGTGGTGA- GTGGGTCCCTAGATGACCATGACCATGTCATGTAATAGCAACATCTATTGAGAAAGG 124398
 QY 63 ----- 63
 Db 124399 GGATGCTGGGTTTTTGAGGGGTGTGAATTTTGTGTAAGGAGGATCATGGGATCAGCCG 124458
 QY 64 ----- LeuProValTyrHis 68
 Db 124459 TGTAGTTGACTTGGGTGAATCATGTGTCCTTCCAGGCATCTGCTGTTTATCAT 124518
 QY 69 SerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPheProProSerHisIlePhe 88
 Db 124519 TCGCTTTTCGTCTGTGGCCACTGAGGACTTCTTGTGCTGGTTTCCCAAGCCACATCTTC 124578
 QY 89 CysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg----- 103
 Db 124579 TGCATAGAGGACGTGGACACTCAAGTCTTGTCTACAGCTACCGTAGGAAGTGAAACCC 124638
 QY 103 ----- 103
 Db 124639 TTGATTTATCTTACAGGTAGTATTCTATTTCAGTCTATAGTAGAGGCTCTTTGAAACCCAGA 124698
 QY 103 ----- 103
 Db 124699 GAGGTTGGGCTGTTGTGTGTAGGCCACACTGCATTCAGTACCCAACTTGGGCTCTGGG 124758
 QY 104 ----- PheTyrPheProAspTrp 109
 Db 124759 AAGCCAGAGCTGAGTGGGCCACCCCATTTCTGAAACCCACAGCTTTTATTTCCCTGACTGG 124818
 QY 110 PheGlyLeuGluThrCysHisArgPheGlyLeuArgLysAspLeuThrSerAlaIleLeu 129
 Db 124819 TTTGGGCTGGAGACATGTACCGCTTTGGGCTGGCGCAAGATTTGACAGGTGCCATCCTT 124878

QY 130 AspLeuHisValLeuGluHisLeuPheAla----- 139
 Db 124879 GACTTACAGCTTTTAGACATCTCTTTGCTCAGTTGGTATTCTTCGGTGGGAAGGCTC 124938
 QY 139 ----- 139
 Db 124939 AACCCAGGATGGGGAGCCCAAGGTAAACAAATCCAGCAGCCAGGACGAAACCTCACCT 124998
 QY 140 ----- GlnHisArgSerAspLeuValSerGlyArgLeuProVa 152
 Db 124999 AACACCCCTTCTGTGGCTGACCCAGCAGCACCGAGTACCTGTTAGTGGGGCCCTCCCGGT 125058
 QY 152 lGlyLeuSerMetLysGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGl 172
 Db 125059 GGGCTTAGCATGAAGGAGCAGGAGAGTTCTTCTGAGCTGGCGGTGCTGACATTGGCCCA 125118
 QY 172 nMetAlaArgGluGlnAlaGlnArgProGlyGluLeuLeuLysThrVal----- 188
 Db 125119 GATGGCTCGTGAAGCAGGCCAGGCCAGGAGAGTCTCTGAAGACGCTCAGGTGAGGACC 125178
 QY 188 ----- 188
 Db 125179 TCGGGCCACCCCTTCGTGGTCTTCTTTGGCTGGGATCTAGGGCCCTATCGTTTCCCTT 125238
 QY 188 ----- 188
 Db 125239 ACTGGAACACATACCTCTTCCCAACTGTGTGACGCCACCTCTCCGTGAGCCTGCCCTG 125298
 QY 188 ----- 188
 Db 125299 CCGAAAGGTCCCTATGCTGTCTGTGTGTCTTCGGGGCCCCCTACCCCTTAAGAGC 125358
 QY 189 ----- SerTyrLysAlaCysLeuProProSerLeuArgAspValI 202
 Db 125359 TCGCGGTGCCCCCCCCACAGTTTACAAAGCTGTCTGCCGCCAGCCTGCGCATGTGA 125418
 QY 202 leGlnGlyGlnAsnPheValThrArgArgArgIleArgThrValValLeuAlaLeuL 222
 Db 125419 TCCAGGCCAGAACTTCTGTGACACGAGCGCATCCGAGACCGTGTCTTGGCGGTGC 125478
 QY 222 euProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLeuAsp 241
 Db 125479 G-CGCTGTGTGCTCCAGCGCCGCTTACCGCTTACGCGCTATGCGCAAGTATATTGGAC 125537
 QY 242 LeuGluArgLeuHisProAlaAlaThrGluThrPheArgValGlyLeuProGlyAla 261
 Db 125538 CTGAGCGGCTTACATCCAGCGGCCACCCAGAGACCTTCCGTGTGGGGCTCCCGGGCGCC 125597
 QY 262 GlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrpSerSerAsn 281
 Db 125598 CAGGAGGAGCCGGGCTTCTGCGTGTGGCGGGGCAACAGCATCTCTCTGGAGCTCCGGG 125657
 QY 282 AspGlu----- 283
 Db 125658 GACCA-GGAGGTGTGGGGCTGGGCTTGAGAGCTGGTCCGCGGGGAGCGTGGCGGT 125716
 QY 283 ----- 283
 Db 125717 TAGTAAAGAGCCCGTCGAACAATCAGGAAAGTGGCTAGTCCGGGCGGGAAGTATTGCA 125776
 QY 283 ----- 283
 Db 125777 CTACGGGCGAGGGGGGGGTCAATTAAAGGTGGTCCGGGGATGGAGCTACCCAGGGA 125836
 QY 283 ----- 283
 Db 125837 CTGACCTTGGGGGGCTGACTAGAAGGGTGAAGTATAGTAGGGGGGGGAGCGACAGAG 125896
 QY 283 ----- 283
 Db 125897 GCGATAGTATGGATTGGATCGGTGGGTGTGACCCGACCCAGGCTGTCTCAGTCCCTTAAC 125956

QY 283 ----- 283
Db 125957 CTCGTCCCGGACCCCATCCACCCCGCTCCCGCGGTAAAGTGGAGGTGGGAAGGAG 126016
QY 283 ----- 283
Db 126017 AAGGGAGGCTGTGCTGGGTGAGGTCTGTAATCCTCTGGGTTCTGCTGTCTCTTTCT 126076
QY 283 ----- 283
Db 126077 CTCCTTTTTTAATGGTTTTGCGAGACAGGGTTTCTCTGTATAGTCTGGCTGTCTGTGA 126136
QY 283 ----- 283
Db 126137 ACTCACTTTGTAGACCAGGCTGGCCTGGAATCTCAGAAATCCGCTCTGCCCTCTCGA 126196
QY 283 ----- 283
Db 126197 GTGCTGTGANTAAAGCGTGGCCACTACACCAGCTACTGCTGTCTCTTTCTTCTCTCAG 126256
QY 284 LeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaPro 303
Db 126257 CTTTTCAGACCTTCTGTGACTTTCGGAATCGTGTATGTACATCAAGCAGGCCCA 126316
QY 304 ArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIle 323
Db 126317 CGTGTGGGTCCGCGCAGGGAAACACCGCTGGTCACCGTCAACGAGATGACGGCCACATC 126376
QY 324 Leu----- 324
Db 126377 GTGTGGGTCTTGGCACCCGTTATTCTGTGTGTGATCTCTGGAAAGGGGTGGTTCGGTT 126436
QY 325 -----GluAlaGluPheProGlyLeuProGluAlaLeuSe 336
Db 126437 GTTGTGAAGATTCTCTGCGCAGGAGCGAGATTTCGGGCTCCCTGAGCGCTGTC 126496
QY 336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCy 356
Db 126497 CTTCTGGCCCTTGTGGATGGGTACTTCGCGCTGATCTCGACTCGCAGCAATTATTCTG 126556
QY 356 sIysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIle 376
Db 126557 CAAGGAGGTGGCGCGCGCGCTGCTGGAGGAGGGGAGCTGTGCCATGTGACCCAT 126616
QY 376 ethr----- 377
Db 126617 CACGTAGGAGCTAGTTGGGTCACAGTGGGATGGGCTGGCAGCGAGGACTGTCTGG 126676
QY 378 -----LeuAspPheAlaIleHisLysLeuLysAlaA 388
Db 126677 GTTGGACCTGGCTGTCTTGGCCTAGGTTAGACTTTGCCATCCACAAGCTGAAGCCG 126736
QY 388 IaGlySerLeuProGlyThrTyrIleLeuArgSerProGlnAspTyrAspSerPheL 408
Db 126737 CTGGCTCCCTCCAGGCACCTACATTTCTCCGCCGAGCGCCGAGGACTATGACAGTTTC 126796
QY 408 euLeuThrAlaCysVal----- 413
Db 126797 TTCCTACCGCTGCGT-CCAGGTTGGTTACTGTGTCGAGGATGGCTTCTCTGGAGA 126855
QY 413 ----- 413
Db 126856 AGTGACCTGAGCTGGGACAGAAAGAGCGTTTCTATTTTTCTTGAAGACACTGGCC 126915
QY 413 ----- 413
Db 126916 TGTAAATCTAGAAATTAAGAGACTGAGGACGAGTCAAGTTTACAGCCCTGCTGAGCTAT 126975
QY 413 ----- 413
Db 126976 ATAGAGAAATATGTTTCTCATGGAGGAAAGGTTTTTTTCCACCAAGCTGAGGGTGT 127035
QY 413 ----- 413

Db 127036 CTTTCATCCATGGCATTTCACAGCATGAGAGGACAGAACTGACACCCATAAATCGCGTCTG 127095
QY 413 ----- 413
Db 127096 ACCTCCACAGTGCACGGTGACTTGCACTTACACAGAAACACACATAAGTGAATGTAATTT 127155
QY 413 ----- 413
Db 127156 TAAGAACTCGAAAGAGCTCAGGAAACAGTCTCAGCATGCGGGGGGGGGGAGAGAAG 127215
QY 413 ----- 413
Db 127216 CATCCTTAGAGTTGGAGACAGNACTGAGCATACAGTCTTCCAGAAAGACGAGAAGGGGT 127275
QY 414 -----GlnThrProLeuGlyP 419
Db 127276 GGGTAGCTCTGCTCCACATCTGGCTCAGGTCCAGGTCTCTTTGACAGCTCCTCTTTGGCC 127335
QY 419 rAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyL 439
Db 127336 CGACTTACAAAGGCTGCCTCATCCGCCAGGACCCACAGCGGGGCTTTCTCCCTGGTTGGCC 127395
QY 439 euSerGlnProHisArgSerLeuArgGluLeuAlaAlaCysTrpAsnSerGlyLeuA 459
Db 127396 TCAGCCAGCCCCACAGAGCCTCGGGAGCTGCTTGCAGCCTGCTGGAATTTCTGGGCTGC 127455
QY 459 rgValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaProArgProLys----- 476
Db 127456 GAGTAGACGTCTGCTCCCTGAACTTAAATCCTGCTGCGCTCCACAGCCCAAGGTGAGT 127515
QY 476 ----- 476
Db 127516 CAGTTCCTCCCTAACCCAGCCCGCAGCTCTTGGGTGGTAGCAGAAACCAAGACCATGG 127575
QY 476 ----- 476
Db 127576 CGATTAGGACAGAAACAGACACTTTTGGCACAGTCTTCAGACATCAAGTCAGTTATTTT 127635
QY 476 ----- 476
Db 127636 CAGTTGGCGCGCGTTTCTTCCCATGGTGAATGCCACTAAGCGCAACCCCTGCTGGACA 127695
QY 476 ----- 476
Db 127696 CATCTCCCTCTGTGGTGTCTGTGTAAACAAAGACCAAGCGCGCTGGAGGAACCCAGT 127755
QY 476 ----- 476
Db 127756 ACGTTCAATGTTCAATCAGCCCCCAAGTCAGAAATCTCAGGAGCAGCTTCCCGGGCGGCC 127815
QY 476 ----- 476
Db 127816 CCAGGCTAAGTATGAGGAGTCGAAATCTTGGAAACCTTGGCTCTGACCTCTGCTTTCTT 127875
QY 477 -----GluLysSerAsnLeuIleValArgArgGlyCysAsnProAlaProAlaP 494
Db 127876 GCCATCCAGAAAGTCCAAATTTGATCGTGGTGGGAAGGGGTGCACCCCGCGCTGCC 127935
QY 494 rGlyCysSerProSerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrA 514
Db 127936 CTGGCTGCTCCCGCTCTGCTGTGGCTGTACACAGCTGAGCTTCCACAAATTCACACGG 127995
QY 514 spSerLeuGluTrp----- 518
Db 127996 ACAGCTGGAGTGGGTGAGGTTCCCGAGGTGAAGCCAGGGCCGGGCCCTTTTGGAGG 128055
QY 519 -----HisGluAsnLeu 522
Db 128056 ACGTGGAAAGCCACTTCAACTGTCTCGCTGGGTGTCTCCCGCCGACGACGAGAACCTG 128115
QY 523 GlyHisGlySerPheThrLysIlePheArgGlySerArgGluValValAspGlyGlu 542

QY 738 lnLeuProAlaLeuLysThrThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrA 758
Db 130334 AGTGGCCGCTCTCAAAATGACAGAACTGCGGGACTTATCACAGTGCATGGCCTATG 130393
QY 758 spProGlyArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrS 778
Db 130394 ATCTGGCGCGGCCCTCTCTCCAGCTATCTCTCAGAGACCTCAACGGCCTCATTACAT 130453
QY 778 ex----- 778
Db 130454 CAGGTCCCGAGAGGGGGGGCGGAGGGCGGAGAGGGTCCCTGAGTGGGTCCAGGAA 130513
QY 778 ----- 778
Db 130514 GTTTGCTGTTCTAGCAGTCAGCGTGAACTAAGACAGCTGTGCAACAACCTAAGACCC 130573
QY 778 ----- 778
Db 130574 GATCTGCAATTTGTTGCACTGTAGCCAGTTGGTGTGTTTACAATGTAGCCCAACTAG 130633
QY 778 ----- 778
Db 130634 CATCTCAGCTGGATTTCTCCCTGCTTACCATCATGGGTGCTGGGATTTGGTTTATTT 130693
QY 778 ----- 778
Db 130694 TGTTATTTCTAAATTTTGTGCAACAGATGTTCTCATCTGAGGGTATCCCAAGTCACAGACT 130753
QY 778 ----- 778
Db 130754 GCGCTACATGCAATGCTGTCTCTCGGGTCCCTTTTGTGTAGCCGTTGGGGTCCCA 130813
QY 778 ----- 778
Db 130814 CAGAGTGTGCTTTGCCCGACAGTGGTCTCTCATTTCCCATTAAGTACTTTGGGGCG 130873
QY 778 ----- 778
Db 130874 TCTCATACGCACAGAGAGCAAGGTTCCAGGTGGAGTTTACACACCAATTTCAGTTTGGG 130933
QY 779 -----AspTyrG 781
Db 130934 ATTGAGGCAGAGACTTACCACTTACCCCACTTCTCTCACGTGATCCCGGCCCTAGATTACG 130993
QY 781 luLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAlaG 801
Db 130994 AGCTCTCTCAGACCCCAACCTGGCATCCCGAGTCTCTCGAGTACGCTGTGC---GGTG 131050
QY 801 lyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLysTyrI 821
Db 131051 GCGCCAGCTCTATGCTGCCAGGACCCCGCCATATTCGAGGAGAGACACCTTAAGTACA 131110
QY 821 leSerLeuGlyLys----- 826
Db 131111 TCTCTTTGCTGGCAAGTGTAGTGGCGGGCATGTGGGGGAGGAACGTGGGTGGAT 131170
QY 827 -----GlyAsnPheGlySerValGluLeu 834
Db 131171 GGGTCAGTGGACACTGCCCTCTCATCTCCACAGGGCACTTTGGCAGCGTGGAGCTG 131230
QY 835 CysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGln 854
Db 131231 TGCCGCTATGACCCCTGGGGGACAATACGGACCCCTGGTGGCAGTGAACAGACTACAG 131290
QY 855 HisSerValProAspGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeu 874
Db 131291 CACACGGGCCACAGACAGAGGAGACTTCCAGCGGAGATTTCAGATCCTTAAGGCTCTG 131350
QY 875 HisSerAspPheIleValLysTyrArgGlyValSerTyrGlyPro----- 889
Db 131351 CACAGCGACTTTCATCGTCAAGTACCGGGAGTCACTATGGGCCAGGTGAGCGGAGCAG 131410
QY 889 ----- 889

Db 131411 CATCTCGGNAACGGGTTGAGTCCCGCTTCTACCCCTTCCAGTAGGGCCCTCTTGAACA 131470
QY 889 ----- 889
Db 131471 AGTCTGTTAACTCCCATGAATCCAGCTTCTATAGCTGGAGGTTGAACGAATCAATACAC 131530
QY 890 -----GlyArgGlnSerLeuArgLeu 896
Db 131531 CCAGTAGAAGGTGCTCGAGAGCTACTGTTCGTGTTCCAGAGTCCGAGAGCCTGCGGTTG 131590
QY 897 ValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly--- 915
Db 131591 GTGATGGAGTACTGCTGCCAGCGGCTGCTTGCAGACTTCTCTGCAGCGCATTCGCGCGCG 131650
QY 916 LeuHisThrAspArgLeuLeuPheAlaTyrGlnIleCysLys----- 930
Db 131651 CTGCACACCGACCGCCTACTGTGTTTCGTTTGCAGATCTGCAAGGTGCGAGCATCGCG 131710
QY 930 ----- 930
Db 131711 AGATCGGATGGGTCTCTGGGGTCCGGATCCCGGGCCACGTGCTATCATGGCTCTCT 131770
QY 931 ----GlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAs 949
Db 131771 GCAGGGCATGGAGTACTTGGTGGCGCGCTGCGGTACACCGTGACCTGGCTGGCGCGAA 131830
QY 949 nIleLeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLe 969
Db 131831 CATCTTGTGGAGAGCAGGCTCATGTGAAGATCGCGACTTCGGCTTCGCTAAGTGTCT 131890
QY 969 uProLeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerProIlePheTyr-- 988
Db 131891 GCCCTGGAAAGGACTACTACGTGTCGCGAGCTGGCCCAAGCCCATCTTTTGGTA 131950
QY 988 ----- 988
Db 131951 CAAGACCAAGACTCCAGAGCTCACGCCACTGAGCCCACTCCCGCACTCTACAGCCTT 132010
QY 988 ----- 988
Db 132011 AAACCGTCCAGACCAGCTCCACCTCAGTAGCTTTGGCTCGCAAAGCCCTTGCCACA 132070
QY 988 ----- 988
Db 132071 GCAGGGTCAGCCCTATAGCCAGATCCAGAAAGCATGCTCAAACTCTACAGATGTACC 132130
QY 988 ----- 988
Db 132131 CCTCTCAACTTAGTCAGAGTACCCCAACCCCAACAGAGTGTCTCCACTCGGTTTA 132190
QY 988 ----- 988
Db 132191 GCCAGCCCCCATTTGTTCTGGCTCCATCTCTTGACCAAGTCTGAAAGCCCGTCACAG 132250
QY 988 ----- 988
Db 132251 GTCTCTTTTCTAGTCCGTCCTCCGAGGGCCCTGCTTTTCAAGCTCTATGGGGTCCGC 132310
QY 988 ----- 988
Db 132311 CCCTCTTAGCGTCTTCTACTTCTCGTCAATCTTACTCTGCTTTTCCCGTCTGCTGCTCT 132370
QY 988 ----- 988
Db 132371 CCCCTTATGACATTTACGCTTATCTGCCAGGCACTTTTCCAGCTGCTGTCATCCAATG 132430
QY 988 ----- 988
Db 132431 CTGACCCACTCAAGCCCGCTCTTAACACTTAGAGTCCCTCTCTAGCTGGCCACGCC 132490
QY 988 ----- 988

Db 132491 CACAGTCCACGAGCCTCTGAGTCGCTTCTGGGGCGGTGCTAGTAGGAAGTCTCC 132550
 QY 988 ----- 988
 Db 132551 TGTTCAGCGCCCTTCTCACCACGAGCTGGCCATATTCACATTGACGACTCCCTCCCC 132610
 QY 988 ----- 988
 Db 132611 AGGCCACTGTTTCACACCTGGCCACCTCTCCTCAGCTGGCCACACACCTCTGTCCGGC 132670
 QY 989 ----- 1006
 Db 132671 TCTCAGTATGCCCCGGAGTCCCTATCTGACACATCTCTCCCGCCAACTCTGACGTGTG 132730
 QY 1006 pSerPheGlyValValLeuTyrGluLeuPheThrTyrCysAspIysSerCysSerProSe 1026
 Db 132731 GAGCTTCGGAGTGTGTGTACGAGCTCTTCACTACTGCGACAAGAGCTGCAGCCCATC 132790
 QY 1026 xAla----- 1027
 Db 132791 CGC-TGTGCGTGGCTGCCCATCCCTCGGTCTCCTCTTGATCTCCAAATCCCTCTCT 132849
 QY 1027 ----- 1027
 Db 132850 GACCTCTAGCCCTATCTGACCCCGCCCTTCTCCTGACCTCCAGATCATCTCTCTGAC 132909
 QY 1027 ----- 1027
 Db 132910 CTCAGTCTCCTCCGGAATCCACGCCCTTCTCCTCATCTCCAGATCCCTTTTGACCC 132969
 QY 1027 ----- 1027
 Db 132970 CAGCCCTTCCCTGACCCCGCCCTTCTCCTCATGAGTCCAGCCCTTCTGACCCCTAAC 133029
 QY 1028 ----- 1037
 Db 133030 AGTCTGCCCTAACTGCTCCTCTCCACAGAGTCTCTGAGCATGATGGGCCCTGAGCG 133089
 QY 1037 gGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluIlyArgArgLeuProPr 1057
 Db 133090 TGAAGGACCCCGCTCTGCGCCCTCTGAGCTCTGAGCTGTCGAGGCGCGAGCCCTCCACC 133149
 QY 1057 oProProThrCysProThrGlu----- 1064
 Db 133150 ACCTCCACCTGCCCCCGAGGTGAGGAGGAGCTCAAGTTCTCAGTTTCAATCT 133209
 QY 1064 ----- 1064
 Db 133210 TGATTGACGAGCGGTGCTGCTGACTGGGAATACTTTACAGGATGAGGAGCGCTCAGCTTC 133269
 QY 1064 ----- 1064
 Db 133270 AAGGGCTTCAGAAATGTGGTTTTCACTTTGTGAGGCAAGGTCTTGTATGAGCCTTG 133329
 QY 1064 ----- 1064
 Db 133330 TCTGCCCTGGAAGTCCGACATCTCTGCTCAGTTCTCCTGACTGAGAGACAAATGT 133389
 QY 1064 ----- 1064
 Db 133390 ATGTGCCACCAACACCCCGCCCACTCCATATTTCAATTTTCTTTTATTATTACTTA 133449
 QY 1064 ----- 1064
 Db 133450 TTTATTTTTCAGTGTGAGATTGAGCCAGGCTTCTACTGATGAGCAAGCAGCG 133509
 QY 1064 ----- 1064
 Db 133510 ACCCTGTCTATGATAATTGCGGGGTGGGGTGGGAGGAGCGCGAGCGCTTCCA 133569
 QY 1064 ----- 1064
 Db 133570 GGAGTGAATGGCGCTATTGGAGAGCCAGGAGAGGTCAATTCCTCCAGTAGGACATGGCT 133629

QY 1064 ----- 1064
 Db 133630 GTGGCTGGGTGGAGTAAGAGAGCAGATGTGCGGAGATGGGGATGTGTGCTGTCT 133689
 QY 1064 ----- 1064
 Db 133690 ACCCTAGCATCTGCAGGCCAGGCGAGAGTGTGCCACAGGCAAACTTAGTTTCTGTT 133749
 QY 1064 ----- 1064
 Db 133750 AGAGATCCATAAGAGGCCATGAGAGCTGTGGAGAGAGTAGGGGCCATTAACTAGT 133809
 QY 1065 -----ValGlnG 1067
 Db 133810 GGAACACACGCTCTCAGCTCAGGCACCCGCCCTTGTCTCCACCACGAGGTTTCAGGA 133869
 QY 1067 uLeuMetGlnLeuCysTrpAlaProGluProHisAspArgProAlaPheAlaThrLeuSe 1087
 Db 133870 GCTCATGCACTGTGCTGGCGCCAGCCGCGCAGCAGCCGCCACCTTCGGCACCCCTGAG 133929
 QY 1087 rProGlnLeuAspProLeuTrpArgGlyArgProGly 1099
 Db 133930 CCCCCAGCTGAGCGCTGTGGCTGGAGAGACCCGGA 133966
 RESULT 11
 LOCUS MMU71201 10326 bp DNA linear ROD 15-JAN-1997
 DEFINITION Mus musculus protein tyrosine kinase JAK3 gene, complete cds.
 ACCESSION U71201
 VERSION U71201.1 GI:1778748
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Musinae; Mus.
 Kumar,A., Toscani,A., Rane,S. and Reddy,E.P.
 Structural organization and chromosomal mapping of JAK3 locus
 Oncogene 13 (9), 2009-2014 (1996)
 97088640
 PUBMED 8934548
 REFERENCE 2 (bases 1 to 10326)
 Kumar,A. and Reddy,E.P.
 Direct Submission
 Title Submitted (18-SEP-1996) Fels Institute, Temple University, 3307 N.
 Broad St, Rm 339 AHB, Philadelphia, PA 19140, USA
 Location/Qualifiers
 1.10326
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="129/J"
 /db_xref="taxon:10090"
 /chromosome="19"
 /map="19p13.1-p13.2"
 join(92..287,441..564,740..851,963..1108,1317..1599,
 2170..2289,2372..2529,2617..2728,3230..3413,3781..3902,
 3998..4129,4596..4680,5022..5149,5642..5774,5905..6056,
 6177..6327,6856..6995,7074..7260,7427..7551,7629..7801,
 8524..8641,8907..9017,9707..10212)
 /product="JAK3"
 join(104..287,441..564,740..851,963..1108,1317..1599,
 2170..2289,2372..2529,2617..2728,3230..3413,3781..3902,
 3998..4129,4596..4680,5022..5149,5642..5774,5905..6056,
 6177..6327,6856..6995,7074..7260,7427..7551,7629..7801,
 8524..8641,8907..9017,9707..9772)
 /function="protein tyrosine kinase"
 /codon_start=1
 /product="JAK3"
 /protein_id="AAB40917.1"
 /db_xref="GI:1778749"
 /translation="MAPSETPLIPORCSLSSEAGALHVLPPRGPPORLUSF
 FGDYLAEDLCVRAAKACILPVYHSLFALATEDSCWPFPSHFICIEDVDQVLVRL

RFYFPMFGLTCHRFRLKDLTSAILDHLVLEHFAQHRSDLVSGRLPVGLSMKEQG
EFLSLAVDLIAQMARQAQRPGLLKTYSKACLPPLSLRDVIOGQNFVRRIRRTTV
LALRVSPARPATAHGOVYDLERLHPAATTFERVLGPAQEEPCLLERVAGDNGIS
WSSGDQLFOTCFDFEIVDVSIKQPTCGSGRHLVTVTRMGHILEAFPGPLPAL
SFVALVDGFRLLICDSRHYCKEVAAPRLLEEABUHGPIITDFAIHLKAAASLP
TYLRRSPQDYSFLUTACVOTPLGPDYKGLIRQDPGAFSLVGLSSPEAKCCLIQ
CWNLSRVDGAALNLTSCCAPRPKESNLIVRRGCTPAPGCSPCCALTLQSFHT
IPTDSLGHENLGHGSETKI PRGRREVVDGTHDSEVLVKVMDSRHNCMESFLEAAS
LMSQVYPHLVKGMEVLSMAGDSIMVOEFVLGALDMVLRKRHLVLSASMKLOVTKOLA
YALNLEDKGLPHGNYSARKVILLAREGGDGNPPFIKLSDPGVSPVTLSEMLDRIPW
VAPECLQEAQTLCLEADKMGFGATTWEVFGSPHAHTLSLEPAKKUFIYEDQQLPALK
WELAGLITOCMAVDGPRPSFRALLDLNGLITSYELLSDDTPGIPSPRDLCCGA
QYACQADPAIFERHLKYLSLKGNGFSGVELCRYPDLONTGPLVAVKLOHSGPQOQ
RUFQRIQILKALHSDFIVKRYGVSPGPGSRURLVMEYLPGLSGDLDFLQRHAAUHT
DLALLFAWOIKGMEVILGARCVHRDLAARNILVSEFAHVKIADGLAKLLPLGXDIY
VVREPGQSPIFWYAPESLSDNIFRSQSDVNSFVLIYELFTYCDKSCSPSAEFRLWNG
PRERGFLRLLELAGRRLLPPPTCTPEVQELMQLCVAPSRITQQPSAP"
BASE COUNT 2088 a 3007 c 2880 g 2300 t 51 others
ORIGIN

Alignment Scores:
Pred. No.: 1,15e-231 Length: 10326
Score: 3935.00 Matches: 1059
Percent Similarity: 32.79% Conservative: 6
Best Local Similarity: 32.60% Mismatches: 31
Query Match: 67.15% Indels: 2163
DB: 10 Gaps: 25

US-09-397-967A-16 (1-1099) x MMU71201 (1-10326)

QY 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
DB 104 ATGGCACTCCAAAGTGAGGAGACACCTTGATCCCTCAGCGCTCTTGACGCTCTCATCC 163
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
DB 164 TCAGAGCGAGAGCCCTGCTGATGCTCTTCTCCCGGGGAGCTGGGCTCCCGACGCA 223
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaIlysAla 60
DB 224 TTGTCAATCTCTTTGGGACTACTTTGGCTGAGGATTTATGTGTGCGAGCTGCCAAGGCC 283
QY 61 Cys----- 61
DB 284 TGTGTGAGTGGTCCCTAGATGACCATGCCATGTAAAGCAACATCTATGAGAAAGGG 343
QY 61 ----- 61
DB 344 ATGCTGGGGTTTGAGGGGTGTGGAATTTTGTGAGGGAGGATCATGGGATCAGCCCGTG 403
QY 62 -----GlyIleLeuProValTyrHisSer 59
DB 404 TAGTTGACTTGGTGAAGTGTGCGCCATCCCTCAGGCATCTCGCTCTTTATCATTCG 463
QY 70 LeuPheAlaLeuAlaThrGluAspPheSerCysTyrPheProProSerHisIlePheCys 89
DB 464 CTTTTCGCTCTGCGCACTGAGGACTTCTTCTGCTGTTTCCCGGAGCCACATCTTCTGC 523
QY 90 IleGluAspValAspThrGlnValLeuValTyrArgLeuArg----- 103
DB 524 ATAGAGGACGTGGACACTCAAGTCTTGTGCTACAGGCTACGCTAGGAGTGAACCCCTTG 583
QY 103 ----- 103
DB 584 ATTTATCTTCAGGTAGTATTTCTATTTCAGTCATAGTAGAGGCTCTTGAAACAGAGAGG 643
QY 103 ----- 103
DB 644 TTGGGCTGTTTCTGTAGGCCACACTGCACTTCAGTGACCCCAACTTGGCGTCTGGAAGC 703
QY 104 -----PheTyrPheProAspTyrPheG1 111
DB 704 CAGAGCTGAGTGGGCCACCCCACTTCTGAACCCACAGCTTTTATTTCCCTGACTGCTTGG 763

QY 111 YLeuGluThrCysHisArgPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLe 131
DB 764 GCTGGAGACATGTACACGCTTTGGGCTGCGCAAGATTTTGACAGTGCCTTCCTGACT 823
QY 131 uHisValLeuGluHisLeuPheAla----- 139
DB 824 ACATGTTTAAACAATCTCTTTGCTCAGGTGTGTATCTTCGGTNGGAAGGCTCAACCCA 883
QY 139 ----- 139
DB 884 GGGATGGGGGACCCCAAGGTAAACAATCCAGCACCAGGACGGAACCTCACCTAACCC 943
QY 140 -----GlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLe 154
DB 944 CTTCTGTGGCCGACCCAGCAGCACCGCAGTGACCTGTGTGAGTGGGCGCTCCCGGTGGGCT 1003
QY 154 uSerMetLysGluGlnGlyPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAl 174
DB 1004 TAGCATGAAGGAGCAGGAGAGTTCCTGAGCCTGGCGCTGTGGACTTGCGCCAGATGGC 1063
QY 174 aArgGluGlnAlaGlnArgProGlyGluLeuLeuLysThrVal----- 188
DB 1064 TCGTGAGCAGGCCAGCGCCAGGAGAGCTGCTGTAAGACGGTCAGGTGAGGGCTCGGNA 1123
QY 188 ----- 188
DB 1124 CCGCTTCGTGTCTTCTTTGGCTGCGGATCTAGGGCCCTATCGTTTTCCCTTCCCGGAA 1183
QY 188 ----- 188
DB 1184 CACATACCTCTTCCCAACTTTGTGAGCCACCCTCTCCGTACGCTGCCCGCGGAAAGGT 1243
QY 188 ----- 188
DB 1244 CCGCTATGTCTGTCTGTGTCTTCGGGCCCCCTCACCACCTTAAGAGCTCGGCNTGNN 1303
QY 189 -----SerTyrLysAlaCysLeuProProSerLeuArgAspValIleGlnGlyG 205
DB 1304 CCCCCCCCCACAGTTTACAAAGCTGTCTCGCGCCAGCCTGGCGGATGTATCCAGGGCC 1363
QY 205 IAsnPheValThrArgArgGlieArgThrValValLeuAlaLeuLeuLeuProCysG 225
DB 1364 AGAATCTGTGACACGAGCGCATCCGAGACCGTGTGTGTGGCGCTGC--GCCTGTG 1421
QY 225 IyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGluArgL 245
DB 1422 GTCGCTGCCAGCGCCGACCGCTACGCTCATGSCCAAGTATAT-CTGGACCTGGAGCGGC 1480
QY 245 euHisProAlaAlaThrGluThrPheArgValGlyLeuProGlyAlaGlnGluLup 265
DB 1481 TACATCCAGCGCCACCCAGAGACTTTCGCTGTGGGCTCCCGGCGGCCAGGAGGAGC 1540
QY 265 roGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrPheSerSerAsnAsp----- 282
DB 1541 CGGGGCTTCTGCTGTGGCGGGGACAACGGCATCTCCTGGAGCTCCGGGAGCCAGGAGG 1600
QY 282 ----- 282
DB 1601 TGCTGGGCTGGCTTGAGAACGGGTCCGCGGGGAGCGCTGGCGGTGTAGTAAAGAAGC 1660
QY 282 ----- 282
DB 1661 CCGTCCGAACAATCAGAAAGTGGCTAGTCCCGGGCGGGCTGATTGCACCTACGGGGCAGG 1720
QY 282 ----- 282
DB 1721 GGGCGGGGTCAATTTAAAGGTGGTCCGGGATGGAGCTGCCCCAGGAGCTGACCTGGG 1780
QY 282 ----- 282
DB 1781 GCGTACTAGAGGTGAGACTAGTAGGGGGCGGGGACGACAGACGCGATAGTAGTATGG 1840

[illegible]

QY 741 laLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGlyA 761
 Db 6214 CTCTCAATGGACAGAACTGGCGGACTTATCACAGTGCATGGCGTATGATCTCGCC 6273
 QY 761 rgArgProSerPheArgAlaIleLeuArgAspLeuAenGlyLeuIleThrSer- 778
 Db 6274 GCGCGCCCTCTCTTCGAGCTATCTCAGAGACCTCAACGGCTCATTTACATCAGTGGCCG 6333
 QY 778 ----- 778
 Db 6334 GAGAGGGGGGGCGGCGAGGGCGAGAGGGTCCCTGAGTGGGTCCAGGAAGTTTGTGG 6393
 QY 778 ----- 778
 Db 6394 TTCTAGCAGTCAGGCTGAATAGACAGCTGCTGCAACAACTTAAGACCCGATCTGCAT 6453
 QY 778 ----- 778
 Db 6454 TTGTTTGCAGTGTAGCCAGGTTGGTGTTTTACAATGTAGCCCAAACTAGCATCTCAGG 6513
 QY 778 ----- 778
 Db 6514 CTGSAITTCNCCCTGCTTACCATCATGCGGTGCTGGGATTGGTTTTTATTTNTTATCT 6573
 QY 778 ----- 778
 Db 6574 TAAATTTTGCACAAGATGTTCTCATCTGAGGTATCGAGTCACAGCGAGTCACTG 6633
 QY 778 ----- 778
 Db 6634 CATGGCTGTCTCTCGGTTCCCTTTTGTCTGTAGCGTTGGATCCACAGAGTGTGC 6693
 QY 778 ----- 778
 Db 6694 CTTTCCCGGACAGTGGTGTCTCATTTCCCAATAGTACTTGTGGGCGTCTCATACGCA 6753
 QY 778 ----- 778
 Db 6754 CAGAGAGCAAGGTTAGGTGGAGTTTACACCAATTCAGTTTGGGATTGAGCGAG 6813
 QY 779 ----- AspTyrGluLeuLeuSerA 785
 Db 6814 ACTTACCACCTTACCCCACTTCTCCTCAGTGTATCCCGCCCTAGATTACGAGCTCTCTCAG 6873
 QY 785 spProThrProGlyIleProSerProArgAspGluLeuCysValAlaGlyAlaGlnLeuT 805
 Db 6874 ACCCCACACCTGGCATCCGAGTCTCAGATGAGCTGTGC---GGTGGCGCCAGCTCT 6930
 QY 805 YrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLysTyrIleSerLeuLeuG 825
 Db 6931 ATGCTGCCAGGACCCCGCCATATTCGAGGAGAGACACCTTAAGTACATCTCTTTGTGG 6990
 QY 825 lYlYs- 826
 Db 6991 GCAAGGTGAGTGGCGGCGCATGTGGGGGAGGAACGTGGTGGTGGATGGGTTCAGGTGA 7050
 QY 827 ----- GlyAsnPheGlySerValIleLeuLeuCysArgTyrAspP 839
 Db 7051 CACTGCCCTCTCATCTCCACAGGGCAACTTTGGCAGCGTGGAGCTGTGCGCTATGACC 7110
 QY 839 roLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProA 859
 Db 7111 CCCTG---GACAAATACGGGACCCCTGGTGGCAGTGAACAGACTACAGCACAGCGGGCCAG 7167
 QY 859 spGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheI 879
 Db 7168 ACCACAGAGGAGACTTCCAGCGGAGATTCAAGATCTTTAAGCTCTGCACAGCGACTTCA 7227
 QY 879 leValLysTyrArgGlyValSerTyrGlyPro- 889
 Db 7228 TCGTCAAGTACCGGGAGTCACTATGGCCAGGTGAGGGCCAGGCAGCANTCTCGGAAACGG 7287
 QY 889 ----- 889

Db 7288 GTTCGAGNNCGCTTCTTACNCTTNNCCAGTAGCCCGGTGAACAGGTCGTTAACTCCCAT 7347
 QY 889 ----- 889
 Db 7348 GAATCCAGCTTCTATAGCTGGAGGTTGAACGAACAATACCCAGTAGAAGTGCTCGAGA 7407
 QY 890 ----- GlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSer 903
 Db 7408 GCTACGTGCTGTTCGAGGTGCCAGAGCTCGGGTTGTGTGATGGAGTACCTGCCGAGC 7467
 QY 904 GlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHisThrAspArgLeuLeu 922
 Db 7468 GGTGCTGCTGAGACTTCTTCGAGCGCATCGCGCGGCCCTGCACACCGACCGCTACTG 7527
 QY 923 LeuPheAlaTrpGlnIle- 928
 Db 7528 CTGTTCTCTTGGCAGATCTCAAGGTGCGACATCGCGAGATCGGGATGGGTCTCTGGG 7587
 QY 929 ----- CysLysGlyMetGluTyrLeuGly 936
 Db 7588 GTCGGATCCCGGGCCACGCTGCTATCATGCTCTCTGCAG-GGCATGGAGTACCTGGT 7646
 QY 937 AlaArgArgCysValHisArgAspLeuAlaAlaArgHisLeuValGlnSerGluAla 956
 Db 7647 GCGCGCGCTGCGTACACCGTGCCTGCGCGCAACATCTTGGTGGAGAGCGAGCT 7706
 QY 957 HisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyr 976
 Db 7707 CATGTGAAGATCGCGACTTTTGGCTCGCTAAGTGTGCTGCCCTGGGAAGACTACTAC 7766
 QY 977 ValValArgGluProGlyGlnSerProIlePheTrpTyrAla- 990
 Db 7767 GTGTCGCGAGCTGCGCAAGGCCCATCTTTTGGTACAA-GACCAAGACTCCAGAGCT 7825
 QY 990 ----- 990
 Db 7826 CACGCCCACTAGCCCCAATCTCCGCACTCTACAGCTTAAACCGGTCCAGACCCAGCTC 7885
 QY 990 ----- 990
 Db 7886 CACCTAGTAGCCCTGCGCTCGCAAGGCTTGCACAGAGGCTCAGCCCTATAGCCAC 7945
 QY 990 ----- 990
 Db 7946 GATCCACGAAGCCATGCTCACAACTCTACAGATGATGCCCTCTCACTTAGTCAGGACTA 8005
 QY 990 ----- 990
 Db 8006 CCCCACCCACCCACAGAGTATGCTCCACNCGGTTTAGCAGGCCCCCAATTGTTCTGGC 8065
 QY 990 ----- 990
 Db 8066 TCCATCTCTTGGACAGTCTGAAAGCCCGTCACAGTCTCTTTCTAGTCCGCTCTC 8125
 QY 990 ----- 990
 Db 8126 GCGAGGCCCTGCTTTCAACGCTCTATGGGGTGTGTCTTAGCGTATTCTACTTCTCGCA 8185
 QY 990 ----- 990
 Db 8186 TCTTACTCTTCTTCTGCTGCTGCTCTCCCTTATGACACCCATGCCCTAGTGTGCC 8245
 QY 990 ----- 990
 Db 8246 AGGCNCTTTCCAGTCTGCTGTCAGTCCGGTGTGACCCACTCAAGCCCGCTCTTAA 8305
 QY 990 ----- 990
 Db 8306 CACTTAGAGTCCCTCTCTCAACCTGCGCCNCGCCACAGGTCCACAGACAGCTCTGAGTG 8365
 QY 990 ----- 990

Db 8366 CCTCTGGGGCCGTGCTAGTAGGAAGTCTCTGTTCACGCCCTTCTCTCACCCACAGAC 8425
QY 990 ----- 990
Db 8426 TGCCCATCTTACATTGACGACTCCCTCCCTGGCCACTGTTTTCACACCCTGCCACTC 8485
QY 991 -----ProGluSerLeuSer 995
Db 8486 TCCTCACTGGCCACACACACCTCTGTCCGCTCTCAGGTATGCCCGGAGTCCCTATCT 8545
QY 996 AspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeu 1015
Db 8546 GACAAATCTTCTCCCGCCAAATCGACGTGTGAGCTTCGGAGTGGTGTGTGACGACTC 8605
QY 1016 PheThrTyrCysAspLysSerCysSerProSerAla----- 1027
Db 8606 TTTCACCTACTGCGACAAAGACTCGACGCCATCGCTGTGCTGCCTGCCCATCCCTCGG 8665
QY 1027 ----- 1027
Db 8666 TCCTCCCTCTTGATCTCCAAATCCCTNCTGACTCTAGNCCCTATCTCTGACCCCGCCC 8725
QY 1027 ----- 1027
Db 8726 TTCCTCCTGACCTCCAGATCACTCCTGACCTCAGTCTCTCCCTCCGGAATCCCAACCTT 8785
QY 1027 ----- 1027
Db 8786 CCTCCTCATCTCAGATCCCTTCTTGACCCAGCCCTTCCCTGANNCCANCCCTTCTCT 8845
QY 1027 ----- 1027
Db 8846 CATGACNTCANCNCCCTTCTGACCCCTAACAGTCTNCCCTAACTTCTCTCTGCCACA 8905
QY 1028 -GluPheLeuAtoMetMetClyProGluArgGluGlyProProLeuCysArgLeuLeuG1 1047
Db 8906 GGAGTTCCTGCGCATGATGGGGCTGAGCGTGAAGGACCCCTCTGCGGCTCTCTGGA 8965
QY 1047 uLeuLeuAlaGluGlyArgLeuProProProThrCysProThrGluVal----- 1065
Db 8966 GCTGTGGCAGAGGGCGCGAGCTCCACACCTCCACCTGCGCCACCAGGT-GAGGA 9024
QY 1065 ----- 1065
Db 9025 AGGAGACTCAAGTTTCTCAGTTTCAATTCTTGATTGACAGCAGCGGCGCTCAGCTGGA 9084
QY 1065 ----- 1065
Db 9085 ATACTTTACAGGATGAGGGAAGCTCAGCTTCAAGGCTTCAGAAATGTGGTTTTCACTTT 9144
QY 1065 ----- 1065
Db 9145 TGTNGAGGCAAGGCTTGTGCTATGTAGCTTGTGTGCGCTTGGAACTCGACATCCTCTCTGCC 9204
QY 1065 ----- 1065
Db 9205 TCAGTTTCTCCTGACTGAGAAGCAAAATGTATGTGCCACCAACACCCCCCACTCC 9264
QY 1065 ----- 1065
Db 9265 ATATTTCAAATTTCTTTTATTATTATTATTATTTTACATTTATTTTACAGTGTGAGATTGAAG 9324
QY 1065 ----- 1065
Db 9325 CCAGGCTTACTCATGCTAGCAAGCAGCACCCTGTGTCATGATAATTCGGGGATGG 9384
QY 1065 ----- 1065
Db 9385 GGGTTGGGAAGGAGCGCGCGTTTCCAGGAGTGAATGGCCCTATTGGAGGAGCCAGG 9444
QY 1065 ----- 1065
Db 9445 AGAGTCAATCCAGTAGGAGACATGGCTGTGGGCTGGGTGGAGTAAGAGAGCATGATG 9504

QY 1065 ----- 1065
Db 9505 TCGNGCAGATGGGGATGTGTGCTGTCACTCCCTAGCCATCTGCGAGGCCAGGAGT 9564
QY 1065 ----- 1065
Db 9565 GTTGCCACAGGCAACTTAGTTTCTGTTCAGAGATCCATAAGGAGCCATGGAGAGCTGTGG 9624
QY 1065 ----- 1065
Db 9625 AGAAGAGTAGNNGCCATTAACTGAATGGAACACCACTCTCAGCTCAGGCATCCGCCA 9684
QY 1066 -----GlnGluLeuMetGlnLeuCysTrpAlaProGluPr 1077
Db 9685 CCTTGATTCCTTCCACACACAGGTTCCAGGAGCTCATGCAGCTGTGGTGGCGCCAG-CC 9743
QY 1077 oHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyAr 1097
Db 9744 GCACGACCGGCCAGCCTTGGCACCTTGAGCCCCAGCTGGACGGCTGTGGCGTGAAG 9803
QY 1097 gPro 1098
Db 9804 ACCC 9807
RESULT 12
AC103459
LOCUS
DEFINITION
AC103459
AC103459.6 GI:30520536
HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus
SOURCE
ORGANISM
Rattus norvegicus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 216847)
Muzny, D. Marie, Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, P.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, Y.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gumaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, J.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, S., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensuhewa, L., Louisegh, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaekemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puazo, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M. A., Reigh, R.,

Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE

Unpublished

REFERENCE

2 (bases 1 to 216847)

AUTHORS

Worley, K. C.

TITLE

Direct Submission

JOURNAL

Submitted (25-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE

3 (bases 1 to 216847)

AUTHORS

Rat Genome Sequencing Consortium.

TITLE

Rat Genome Sequencing

JOURNAL

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 10, 2003 this sequence version replaced gi:24953141. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GHQO

Center clone name: CH230-30H1

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 199697 bases at least Q40

Consensus quality: 201698 bases at least Q30

Consensus quality: 203226 bases at least Q20

Estimated insert size: 208186; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* the accession number will be preserved.

1. 216847: contig of 216847 bp in length.

Location/Qualifiers

1. 216847

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

FEATURES

source

/db_xref="taxon:10116"
/clone="CH230-30H1"

misc_feature

85405..86754

misc_feature

/note="wgs contig"

208102..208763

misc_feature

/note="clone boundary"

clone end:T7

misc_feature

site: EcoRI

end sequence: BH291939"

215623..216847

misc_feature

/note="wgs end extension"

clone end:T7"

BASE COUNT

47554 a 54739 c 54387 g 48109 t 12058 others

ORIGIN

US-09-397-967A-16 (1-1099) x AC103459 (1-216847)

Qy 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20

Db 123069 ATGSCACCTCCAGCGAGGAGACACCTCTCATCTCTCAGCGCTCTTGTAGCCTCTCATCC 123128

Qy 21 SerGluAlaGlyAlaLeuHisValLeuLeuProGlnArgGlyProGlnArg 40

Db 123129 TCAGAGCAGGAGCGCCCTGCATGTCTCTCTCCCGGGACCTGGCGCTCCCCAGCGA 123188

Qy 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60

Db 123189 CTGTCATCTCTTTTGGGACTACTTGGCTGAGACCTGTGTGCGGGCTGCCAAGGCC 123248

Qy 61 CysGlyLeu 63

Db 123249 TGTGTGA-GTGGTCTCTAGTAGACGAGGCCACTTAGTAGGCTGAGCAACATCTTGAG 123307

Qy 63 63

Db 123308 AAAAGGAGTGTGGGGCTTTGCGGGATGCGTAGAAGTTTGTAAAGGAGGATCATGGCT 123367

Qy 63 63

Db 123368 CAGCTGTGTGTTCATTGGTGTAAGTGGGTCATCTCTTTCAGGCATC 123427

Qy 64 LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPhePro 83

Db 123428 CTGCCCGTTTATCATTCGCTCTCGCTCTGGCCACCGAGGACTTGTCTTGTGTGTCTCC 123487

Qy 84 ProSerHisIlePheCysIleGluAspValPheThrGlnValLeuValTyrArgLeuArg 103

Db 123488 CCAAGCCACATCTTCTCCATAGAGACGTGGACACTCAAGTCTTGGTCTACAGGCTACGG 123547

Qy 103 103

Db 123548 TAGNAGTGAGCCCTCTATTACCTTCAGGGTAGTATTCTTCTCTTTTCAAAGAT 123607

Qy 103 103

Db 123608 TTGTTTATTATTATATATAAATCAATTGTTGCTGTCTTCAGACACACAGAGAGGCGA 123667

Qy 103 103

Db 123668 TCAGATCTCATTTATAGATGGTTAAGCCACCATGTGTTATGGGATTTGAACCTCAGGACC 123727

Qy 103 103

Db 123728 TCTGGAAGAGCAATCAGTGTCTCTCACTGCCATCTCTCCAGCCCCCAGAGTAGTATTCTTCT 123787

QY	617	SerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeu	-----	633
DB	128164	AGCTGAAACTGCAGGTGACCAAGCAGCTGGCTATGCCCTTAACCTACT		128222
QY	633	-----	-----	633
DB	128223	TCCTGTGTCGTTGTTTATTTCCTCTACCAATGAGGAAATGAAGAGTAGGCTCTGTAGTTT	-----	128282
QY	633	-----	-----	633
DB	128283	GTATAGGAGCTCGGTGAGGACTGAAGCAAGCAAGGATGTGAGCCTGTTCAGGGGGTTG	-----	128342
QY	633	-----	-----	633
DB	128343	AGAGCATCAAGGACTGTAAACCCATCGGGTGTCCGAGCCAGGCTGTGGGGACAGGGCATA	-----	128402
QY	633	-----	-----	633
DB	128403	GAGATGGAAACTGGTTAAGAGGAGAGGGTGTCTTGAATGGTGAATAGAGGCTTTAACTG	-----	128462
QY	633	-----	-----	633
DB	128463	TCCCTAAAGCCATGGAAAGGTAGAGAAAGGGCGTGGTGGTTAAGGGTAGGTATGATTAG	-----	128522
QY	633	-----	-----	633
DB	128523	TGTAAGGCGAGTTAGGGTCAACAGACATGTTCTCATGGACTTACAGTGACTGTCAAAGC	-----	128582
QY	633	-----	-----	633
DB	128583	CTGTTCTGTCGGGAGCATCAGGTAGGCTTGAAGGATGCAGAGAGTTTGTCTTAATAG	-----	128642
QY	634	-----	-----GluAspLys	636
DB	128643	TGTTCAATTCATAAACAACCTCATCCACATTCCTCTGTTCTCGAACAACCTTAGGAGACAAA		128702
QY	637	GlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGlyAsp	-----	656
DB	128703	GGCTCCCTCATGGCAACGTCTCAGCGAGGAAGGTGCTGCTAGCTCGTGAGGGGTTGAT	-----	128762
QY	657	GlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeu	-----	676
DB	128763	GGGAATCCACCTTTTCATCAAGCTGAGTGATCTCTGTTGTCAGCCCAACTGTTTGTAGCCTG	-----	128822
QY	677	Glu	-----	677
DB	128823	GAAAGTGAGTAGCCGAGCCTCACCCAACCTATACCTGTGTCTGTCTGCACACACA	-----	128882
QY	677	-----	-----	677
DB	128883	GTTACAGGGCGCTGCATGTCGCAAGGAGACCCGCTGTGCAGGAGTGACCGTGACCATT	-----	128942
QY	678	-----MetLeuThrAspArgIleProTrpValAlaProGluCysLe	-----	691
DB	128943	CTCACTCCACATCCTCCAGTGCTCACTGACAGGATTCCTTGGGTGGCCCGGAGTGCT	-----	129002
QY	691	uGlnGluAlaGlnThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyValAlaThrThrTr	-----	711
DB	129003	CCAGAGCTGGAACGCTCACTTGGAGGCTGCACAGTGGGGCTTTGGCGCCACACG	-----	129062
QY	711	pGluValPheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaLys	-----	728
DB	129063	GGAGGTGTTACGCGGGGACCCCATGCACATTACCTCGCTGGAGCCCGCCAA-GGTCAGAG	-----	129121
QY	728	-----	-----	728
DB	129122	CAGTCTCAGCCTTATGGATCTGGTGTGGTAGGGTTAAACCACGTCAGCGGCAT	-----	129181
QY	729	-----LysLeuLysPheTyrGluAspGlnGlyG	-----	738
DB	129182	CTCAAAACCCAAACGGAACACTCGGCCCTCAGAACGCTGAAGTTCTATGAGGACCGGGACA	-----	129241
QY	738	nLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAs	-----	758

D	b	129242	GCTGCGGCCCTCAAAATGGACAGAGCTGGAGGGGCTCATCGCACAGTGCATGGCGTATGA	129301
Q	y	758	pProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAunSgNyleThrSe	778
D	b	129302	TCTGGCCCGGCCCTCCTTCGAGACCATCCTCGAGACCTCAACGGGCTCATTTACATC	129361
Q	y	778	I-----	778
D	b	129362	-AGGTACCTGGGTGGAGAGGTCGCCGAGCTGGGTCCGGAAGTTGCTGGCTCCAG	129420
Q	y	778	-----	778
D	b	129421	AGTCAATGTAGTAAGACAGACTGTGTGAATAAACCTGTGGCTAAGATAAAAACCTCAATCA	129480
Q	y	778	-----	778
D	b	129481	GCATTGTTTTGCAATGTAGTCCAGCTAGTGTTTTACAGTAGCCCCAGGTGGCCTC	129540
Q	y	778	-----	778
D	b	129541	CAGACTGGATTCTCTCTGCCTCATCGTCTCGATTGGTTTTATTGTTATCCGGGTC	129600
Q	y	778	-----	778
D	b	129601	ACAGCTGTGTCACTGTATGGCTCTGGCTCTCGGGTCCCCTTGTGTCTGTAGCCCTTG	129660
Q	y	778	-----	778
D	b	129661	GGATCCACAGAGTGTGCTTTGCCGGACTGTGATCGTTCACACTCCCACAAGTACC	129720
Q	y	778	-----	778
D	b	129721	CGTGGGCATCTCATTAATGCACAGAGAGCAAGTTTCAGATGGAGTTCACACCAAATTC	129780
Q	y	779	-----AspTy 780 	
D	b	129781	AGTTTGGGTTGAGSCACAGAAAGATACCCCACTTACTCATGTGACCCCGCCCTAGATT	129840
Q	y	780	rGluleuLeuSerApproThrProGlyileProSerProArgaspGluLeuCysValal	800
D	b	129841	CGAGCTCTCTCAGACCCCAACACCTGGCATCCGAATCCCGAGATGAGCTGTGI--GG	129897
Q	y	800	aGlyAlaGlnLeuTyraLcysGlnAspProAlaIlePheGluArGHisLeuLyety	820
D	b	129898	TGGTCCCAGACTCTATGCTGCCAGGACCCGCCATCTTCGAGGAGACACCTTAAGTA	129957
Q	y	820	rIleSerLeuLeuGlyLys-----	826
D	b	129958	CATCTCCCTGTGGGCAA-GGTGAGTGAGCGAGCTTGTGAGCGAGGCTTAAGTTGGCG	130016
Q	y	827	-----GlyAsnPhEGlySerVa 832	
D	b	130017	GGTGGAGGGTCAGSCAGAAACTCGGCTCTCTTCCTCCACAGGGGAACCTTCGGCAGCGT	130076
Q	y	832	lGluleuCysArgTyraSpProLeuGlyAspAsnThrGlyProLeuValAlavalLysGl	852
D	b	130077	GGAGCTGTCCCGCTACGACCTCTGGGAGACAATACGGGACCCCTGGTAGCAGTGAAGCA	130136
Q	y	852	nLeuGlnHisSerValProaspGlnGinArgAspPheGlnArgGluIleGlnIleLeuLy	872
D	b	130137	GTTACAGCACAGCGGCCAGAACACAGAGGGGACTTCACGCGAGAGATTGAGTTCCTCAA	130196
Q	y	872	sAlaLeuHisSerAspPheIleVallysTyraGlyValserTyraGlyPro-----	889
D	b	130197	GGCTCTACACTGACTTCACTCGTCAAGTACCGGGAGTACGCTATGGGCTGCTGAGAG	130256
Q	y	889	-----	889
D	b	130257	CGGCGACGACTCGGGAAGGGTTTGTAGTCCCGCTTCTACCTCTTCCAGTTGGGGAGTG	130316
Q	y	889	-----	889

130317	TTGAGCAAGCGCGTTAAACCCCATGAATCTCAGCTTCATATAGCTGGGGCTTTGACGGAACC	130376	QY	988	-----	988
989	-----GlyArgGlnSerLe	894	QY	988	-----	988
130377	AATGCACCCGGTAAAGGTGCTCAGAGACTACTTTCCCTATCCGAGAGTCCGACAGAGCCT	130436	Db	131457	TGCAGTCTTGCTTCTAGTCTCCTGACACACACCCCTTTCTAGTCCCCCAACCTTAACC	131516
989	-----TyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrps	1007	QY	988	-----	988
130437	CGGTTAGTATGAGTACCTGCCAGCGGCTGCTGGGACATTCCTGACAGCGCCACCG	130496	Db	131517	ACACCTGTCCAGATCCTTACTGTGGCCACACCCCTTTTCTGACTAGGCCTCTGTCCCGCTC	131576
914	gGly---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLys	930	QY	989	-----	1007
130497	CGCGCGCTGCACACGACCGCTACTCTGTTCCCTGGCAGATCTGCAAGGTGCGAGC	130556	Db	131577	TCAGGTAGCCCTGAGTCCCTATCTGACAACATCTTCTCCGCAATCTGACGTGTGGA	131636
930	-----	930	QY	1007	erPheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerA	1027
130557	AGCGTGGAGATCCGAGATCGGGATTCCTGGGGTCCGTGATCCCGGGCCACATGCTATCACG	130616	Db	131637	GCTTCGGCGGTCTGCTATGAGCTCTTACCTACAGCGACAAGAGTGCAGCCCATCCA	131696
931	-----GlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaA	947	QY	1027	la	1027
130617	GCCCCCTACAGGGCATGGAGTACCTGGCGCAGCGCTGCTGTCATCGTACCTGGCTG	130676	Db	131697	C-TGTGCGTGCAGAGTCCCATCTCTCCCTCAGTCTCTTCTGTGACCTCTAACCCCTCCTTT	131755
947	laArgAsnIleLeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaL	967	QY	1027	-----	1027
130677	CGCGCAACATCTGCTGGTGGAGAGTGAAGGCGCATGTGAAGATTGCTGACTTCGCGCTCGCTA	130736	Db	131756	GACCTCCAAACTTTTCTTCTTGTCTTCCAGCCTTCTCTGACCTCAGCCCTTCTTCCCA	131815
967	ysLeuLeuProLeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerProIleP	987	QY	1027	-----	1027
130737	AGCTGCTGCCTCTCGAAAGAGACTACTAGTGGTCCGAGCGCGGCCCAAGTCCCATCT	130796	Db	131816	ACCTCCAGTCTCTCTCCTGACCTCCAGACCACCTTCTTGACCTCCAAACCTTCCTCTCTGA	131875
987	heTrp-----	988	QY	1027	-----	1027
130797	TCTGTTATAGACCAAAACCCCTGGTGGGATTTAGTCACTGCTAGAGCGCTTGCT	130856	Db	131876	CCTCCAGTCCCTTCTCCTGACTTCCAGCCCTGTCTCCTGAGCCCGAGCCCTCCTTCTCTGA	131935
988	-----	988	QY	1027	-----	1027
130857	AGGAAGCGCAAGGCCCTGGGTTGCGTCCCGAGCTCCGMAAAAGAAAAA	130916	Db	131936	GTCACAGCTCTTTTAAAAAGATTATTCAITTTATATATATATAGTACACTGTAGCT	131995
988	-----	988	QY	1027	-----	1027
130917	GACCAAAACCCCTGACGCCACGCCACTGAGCCCCAGACTCCACACTTTACTGCTCAA	130976	Db	131996	GTCITTCAGATACACAGAGAGGGCATCGGATCTCTTTACAGATGATTGTGAGCCACCAT	132055
988	-----	988	QY	1027	-----	1027
130977	ATCGGTCCACTTCTCAAGCTTGGCCCTCGCAGAGGCGCGCCACAGAGGGGTGAGCC	131036	Db	132056	GTGGTTGTGGAAATGGAACCTCAGGACCTCTGGAAGAGCAGTCGGGGCTCTTAACCGATG	132115
988	-----	988	QY	1027	-----	1027
131037	GTAGCCACGCCACGAGGCCACGCCACAGCTTTATPACCAGTACCAGCTAACTGTGTA	131096	Db	132116	GGCCATCTCTCCAGCCCGAGTCCAGCTCTTTTTCCTCACCCTCCAGATCCCTCTCTGAC	132175
988	-----	988	QY	1027	-----	1027
131097	GTCCGGACTACCGACCTCTCCACAGAGTATGGTTAACTCATGTTCTGGCTTCGC	131156	Db	132176	CCCCACCTGTCTCCTGACCCCTTCAGGCCTTCTTATGATATCAGCCCTTCTCTGACCCCTTCTCTGACCC	132235
988	-----	988	QY	1028	-----GluPheLeuArgMetM	1033
131157	CCTCAGTGACCACAGTCAGCAAGCCCTTACAGAGTCTCTTCTAGCTCCGCGCCCTCCGCA	131216	Db	132236	CAGCCCTCCCTCTCTGTCGCTTAACCTGTCTCTCCACAGAGAGTCTCTGCGCATGA	132295
988	-----	988	QY	1033	etGlyProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyA	1053
131217	CGGCGCGGCTTTTACGCTCCATGGGGTTGCGCCCTCTAACGTCACCCACTTCTGTCTAT	131276	Db	132296	TGGGCGCTGAGGTGAGGGATCCCGCTCTGCCACTCTTGGAGCTGTGGCAGAGGGCC	132355
988	-----	988	QY	1053	tgArgLeuProProProProThrCysProThrGlu	1064
131277	CTTCTCCGCTTCCCGTCTGCTGCTCTCTCTATGGAATTATACGTACTGTGATGC	131336	Db	132356	GGCGTCTCCACACCCCTCCACCTGCTTACTGCTGAGGTGAGGAGACAGACTCAAGTTTCC	132415
988	-----	988	QY	1064	-----	1064
131337	CTATCTGCCCGCCAGCACTTTCCCGATTGCATCATCCAAAGCCCGTCCAAATCCACTTA	131396	Db	132416	CAGTTTCCATTCCTGATTGACAGCATTTGGTCAATGACTGGGAAGGCTTTTACATGAGAGGC	132475
988	-----	988	QY	1064	-----	1064
131397	GAGTCTCACCCCTCAACCTGGCCAGCCCGCCACAGTCCCGACAGACGCTCTGAGCGCGCTTA	131456	Db	132476	TCAGCTTCAAGGGCTTCAGAAATGTGGTTTTTTCACTTGTGTGAGGCAAAAGTCTTGTCTATG	132535


```
QY 1064 ----- 1064
Db 132536 TAGCCTCTCTAGCCTGGAACGACATCTCTGCTGCTCAGTCTCTTCTGAGCTGAGAAG 132595
QY 1064 ----- 1064
Db 132596 ACAATGCGAGTGCCACCAACCACTCTCACTCCACATTTTGTGTTTCTTTATTAT 132655
QY 1064 ----- 1064
Db 132656 TTATTTTTTACAGTGTGAGATTGAAGGACATGCTAGGCAAGCACACATCTATGTCA 132715
QY 1064 ----- 1064
Db 132716 TGACAATGAAGGGGCTCGCAGGCCACTCCCGGGCAATAGTGATCTTTTGGAGGAGCC 132775
QY 1064 ----- 1064
Db 132776 ACTAGAGTTCAATCCAGTAGGAAGCAGAGCTTAAGAGAGAACAATGTGTTGCAAGGTG 132835
QY 1064 ----- 1064
Db 132836 GCGAATGTGTGCCGTCACTCTAGCCATCAGCAGGCCAGGCGAGGAATACCAAGCC 132895
QY 1064 ----- 1064
Db 132896 AAAGTTAGTTTCTGTTTACAGATCCATAGGAGCCAGGAGCTATGAGCAGTGGGT 132955
QY 1064 ----- 1064
Db 132956 GACTTCGAATGTGTGCTTGAAGTGAAGTGAAGAACCAAGCTCAAGCAGCCGCC 133015
QY 1065 -----ValGlnGluLeuMetGlnLeuCysTrpAlaProGluPro 1077
Db 133016 CACCTTGCTTCCCACTCAGTTCAGAGCTCAGTGTGCTGCTGCGCCCAACCCA 133075
QY 1078 HisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArg 1097
Db 133076 CAAGACCGGCCAGCATTTGACACCTTGAGCCCCAGCTGATGCGTGTGGCGCGGAGCC 133135
QY 1098 ProGly 1099
Db 133136 CCCGGA 133141

RESULT 13
AF034576
LOCUS AF034576 3644 bp mRNA linear VRT 03-SEP-1998
DEFINITION Gallus gallus Janus tyrosine Kinase (JAK) mRNA, complete cds.
ACCESSION AF034576
VERSION AF034576.1 GI:2645986
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus;
Sofer, L., Kampa, D. and Burnside, J.
TITLE Molecular cloning of a chicken JAK homolog from activated T cells
JOURNAL Gene 215 (1), 29-36 (1998)
MEDLINE 98332716
PUBMED 9666067
REFERENCE 2 (bases 1 to 3644)
AUTHORS Sofer, L., Kampa, D. and Burnside, J.
TITLE Direct Submission
JOURNAL Submitted (14-NOV-1997) Animal Science, University of Delaware, 40
Townsend Hall, Newark, DE 19716, USA
LOCATION/Qualifiers
FEATURES
source
1. 3644
/organism="Gallus gallus"
/mol_type="mRNA"
/sub_species="domesticus"
/db_xref="taxon:9031"
```

```
/cell_type="lymphocyte"
/tissue_type="spleen"
/notes="Cells activated with concanavalin A"
1. 3644
/gene="JAK"
91. 3411
/gene="JAK"
/notes="similar to JAK protein kinase"
/codon_start=1
/product="Janus tyrosine kinase"
/protein_id="AAC34195.1"
/db_xref="GI:2645987"
/translation="MAPIGEEETPLIGERSCSLSSEPEPTLQVLYHRGHAPDPSAAT
LIFTGEYTABELCVHAAKGLPICHPLFALATEDLSWYPNNHFTVEDARSQV
VRIIRFPFNWCGQGVHVRPQLSPDRSPVLDYPIVLYLFAQSRSDFIAGRMELALS
AQBEELSLAVLMLRIAERKQSPKVFVYKTCIPQLRCQIQOHSFLTRKRIR
RVAQSLRMGRCVHDGCLLKLILDLERLCCWAEESFAHGPADIAIHVSTDG
SVASCVGSSRHFCDPIADVSIKOASRDGVPVENRIVTLTKTNRVLEVPFTR
EALSFWALVDGYRLTADAHYFCKEAPPRLLEDMEQCHGPISSEFAVRKLKAGS
HPGLYLRSPDPFSLTLTCAETRSQDYKRLIRDEDEDFWLAGLARFCSLQE
LUGTGGCCGQAEGAHLLDTCPPPREKSNLLIVRSGCPRPSPAPRSPNQMS
FKIDPESLIRGESLQGSFTHIYKIKRDQKDEFYQTPVLYKVMDSHRNCSFSL
EASIMSQLSHKHLVLLHGVSLGDSIMVQYIRHGPLDLYLKKNHSEKVTTSKLO
VAKQAYALNYLEDKKITHGNVSAKVVLTREGDAASSPPFIKLPDGVSTVLAK
WLVRIPWAVECLSDPSLALPADKMGFGATLWEIFSGGNMPVSLLEPQKLQFVDS
RLQLPAPRSELAALIACTIYAPRRPCFALIRINDISSDIYELLSLSPDVTLR
ESCQYEHVAGHGAQFEERHLKYISLLGKGNFGSVELCRYPDLGDSGTGELVAVKL
QDSAKELQDFEREIQLHSLQHDFIVRYGVYSGRRRLRVNMEYLPDGLCLRYLO
QNRHLEHRTLLLYAWQICKMEYLGACORCVHRDLASRLNIVELSTHVKIGDFGLAKL
LPQDKDYVVOEPGOSVFWAPESLADNVFSRASDVMSFGVLVLELFTYSNKSRSFS
EEFLHMGPEKPAQIICHLLELLKDSRLRPVPPGCPMEVYAMMLSCWAFAPASRPFTT
ELIARVEALRDGRGTAGH"
BASE COUNT 697 a 1173 c 1150 g 624 t
ORIGIN
Alignment Scores:
Pred. No.: 1,92e-195 Length: 3644
Score: 3336.00 Matches: 670
Percent Similarity: 72.89% Conservative: 142
Best Local Similarity: 60.14% Mismatches: 272
Query Match: 56.93% Indels: 30
DB: 5 Gaps: 14
US-09-397-967A-16 (1-1099) x AF034576 (1-3644)
QY 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
Db 91 ATGGCCCGCTGGCGGAGGAGACGCCGCTGATCGTGAGCGCTCTCTGAGTCTGTCTTCC 150
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyPro---GlyProProGln 39
Db 151 TCGGAGCCCGGACGCTGAGGTGTACCTTACCACCGGGGAGCCCATGGCGCCCGCCGAC 210
QY 40 -----ArgLeuSerSerPheGlyAspTyrLeuAlaGluAspLeuCysValArg 56
Db 211 TCCGCTGCCACCTTACCTTTCGCGAGTACACGCGCGAGGAGCTGTGGGTGCAC 270
QY 57 AlalAlaLysAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGlu 76
Db 271 GCTGCAAGGCTCGGAGTGTCTGCCCATCTGCCACACCCCTCTTCGCGTTCGCTACGGAG 330
QY 77 AspPheSerCysTrpPheProSerHisIlePheCysIleGluAspValAspThrGln 96
Db 331 GATCTCAGTGTCTGTGATCCCCCCCCCAACCATCTTACGGTGGAGGACGCCCGCAGCCAG 390
QY 97 ValLeuValTyrArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHis 116
Db 391 GTGGTGGTATAGGATCAGGTCTTCTTCCCAACTGGTGTGGGCGAGGCGGCGAGTGCAC 450
QY 117 ArgPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHis 136
Db 451 CGCTTCCAGTGCACCAAGTGACCGCCCGCCCGCTGTGGACTACCTGTGATCGATTAC 510
```



```

Db      2635 GAGCTGGTGGCCGTGAGAGAGCTGCACAGGATTCGCCAAGGAGCTGCAGGACTTTGAG 2694
QY      866 ArgGluLeuGlnLeuLeuLeuAlaLeuHisSerAspPheLeuValLysTyrArgGluVal 885
Db      2695 AGGAGATCCAGATCTGCTGCTGCACACGACTTCATGTCAGATACCGGGGGCTC 2754
QY      886 SerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCys 905
Db      2755 TGCTACAGCGCTGGGGCGCGGGGCTGCGGTGGTGATGAGTATCTGCTGACGGCTGC 2814
QY      906 LeuArgAspLeuLeuGlnArg-----HisArgGlyLeuHisThrAspArgLeuLeuLeu 923
Db      2815 CTGAGGAGTACCTGCAAAAGACACAGCACCGC---CTGGAGCACCGCAGCTGCTGCTC 2871
QY      924 PheAlaTrpGlnLeuLeuGlyValMetGluTyrLeuGlyAlaArgArgCysValHisArg 943
Db      2872 TAGCCCTGGCAGATCTGCAGGAGCATGGAGTACTCGGGGCGCAGCGCTCGTGGCACCGC 2931
QY      944 AspLeuAlaAlaArgAsnLeuLeuValGluSerGluAlaHisValLysLeuAlaAspPhe 963
Db      2932 GACTTGGCCAGCAGGAACATCTGCTGGAGAGCGAGACCCACGCTCAAGATCGGTGACTTC 2991
QY      964 GlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArgGluProGlyGln 983
Db      2992 GGCTGGCCCAAGCTGCTGCCACAGGACAAAGGACTACTACGTGTGTGAGGAGCCCGGCGAG 3051
QY      984 SerProilePheTrpTyrAlaProGluSerLeuSerAspAsnLeuPheSerArgGlnSer 1003
Db      3052 AGCCCGCTGTTCTGTACCCACCGGAGTCTGCTGCTGACACGCTCTTCCCGAGCATCC 3111
QY      1004 AspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCys 1023
Db      3112 GACGTGTGAGGCTCGGGGTGCTGCTATGAGCTCTTCCACCTACAGCAATAAGACAGG 3171
QY      1024 SerProSerAlaGluPheLeuArgMetMetGlyProGluArgGluGlyProProLeuCys 1043
Db      3172 ACCCCCTCGGAGAGTCTTCCACATGATGGGCCCTCGAGAAACCCACAGCATCTCTGC 3231
QY      1044 ArgLeuLeuGluLeuAlaGluArgArgLeuProProProProProProProProPro 1063
Db      3232 CATTTGCTGGAGCTCTGAGGAGCTCCCGGGGCTCCCGGGTCCCGGGTGGCCCATG 3291
QY      1064 GluValGlnGluLeuMetGlnLeuCysTrpAlaProGluProHisAspArgProAlaPhe 1083
Db      3292 GAGGTGTATGCGATGATGCTGAGCTGCTGGGCTGCTGGCCCGGAGTCCAGACCCACCTTC 3351
QY      1084 AlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArg 1097
Db      3352 ACCGAGTTGGCAGCCAGGGTTGAGGCACTGCGGGGATGGCCGT 3393

RESULT 14
AX203084
LOCUS      2914 bp      DNA      linear      PAT 30-AUG-2001
DEFINITION      Sequence 5 from Patent WO0152892.
ACCESSION      AX203084
VERSION      AX203084.1  GI:15392441
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE
AUTHORS      Vasios,G.
TITLE      Jak/stat pathway inhibitors and the uses thereof
JOURNAL      Patent: WO 0152892-A 5 26-JUL-2001;
GENZYME      CORPORATION (US)
FEATURES
source
1..2914
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT      566 a 907 c 809 g 632 t
ORIGIN

```

Alignment Scores:

```

Pred. No.:      5.8e-189      Length:      2914
Score:          3228.50      Matches:      614
Percent Similarity: 87.63%      Conservative: 45
Best Local Similarity: 81.65%      Mismatches: 90
Query Match:    55.09%      Indels:      3
DB:             6              Gaps:      3

```

US-09-397-967A-16 (1-1099) x AX203084 (1-2914)

```

QY      350 AspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGluGluAla 369
Db      5 GACTCCACGACTTCTTTCAGAGGAGTGGACCCGAGGCTGCTGGAGAGTGGCC 64
QY      370 AspValCysHisGlyProLeuThrLeuAspPheAlaHisLysLysLeuLysAlaAlaGly 389
Db      65 GAGCAGTGCACCGGCCCATCACTCTGGACTTTGCCATCAACAAGCTCAAGACTGGGGC 124
QY      390 SerLeuProGlyThrTyrLeuArgArgSerProGlnAspTyrAspSerPheLeuLeu 409
Db      125 TCACGTCTGGGTCTCTATGTTCTCCGCGCAGCCCGCAGACTTTTCACAGCTTCTCTCCTC 184
QY      410 ThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCysLeuLeuArgGlnAsp 429
Db      185 ACTGCTGTGTCAGAACCCCTTGTCTCTGATTATAAGGGCTGCTCATCGGGCGCAGC 244
QY      430 ProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeu 449
Db      245 CCCACAGGAACCTTCTTCTGGTGGCTCAGCCGACCCACACAGCACTTCTCGAGAGCTC 304
QY      450 LeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSer 469
Db      305 CTGGCAACCTGCTGGAGTGGGGGCTGACGTAGATGGGTGGCGAGTACCCTCACTTCG 364
QY      470 CysCysAlaProArgProLysGluLysSerAsnLeuLeuValValArgArgGlyCysAsn 489
Db      365 TGCTGTATCCCGACAGCCCAAGAAAGTCCAACTGATCGTGTCCAGAGAGTTCACAGC 424
QY      490 ProAlaProAlaProGlyCysSerPro---SerCysCysAlaLeuThrGlnLeuSerPhe 508
Db      425 CCACCCACATCATCTTGGTTTCAGCCCAATCCCAATACCAGCTGAGTCAGATGACATT 484
QY      509 HistThrLeuProThrAspSerLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThr 528
Db      485 CACAAAGATCCCTGCTGACAGCCTGGAGTGGCATGAGAACCTGGGCCATCGGTCCTTACC 544
QY      529 LysIlePheArgGlySerArgArgGluValValAspGlyGluThrHisAspSerGluVal 548
Db      545 AAGATTACCGGGGCTGTCGCCATGAGGTGGTGGATGGGGAGGCCGAAAGACAGAGGTG 604
QY      549 LeuLeuLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaAla 568
Db      605 CTGCTGAAGGTTCATGGATGTCACCAAGCACCAAGAACTCATGAGTGCATTTCTGGAACAGCG 664
QY      569 SerLeuMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAla 588
Db      665 AGCTTGATGAGCAAGTGTCTGACCGGCATCTCGTGTGCTGCCAGGGGTGGCATGGCT 724
QY      589 GlyAspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArg 608
Db      725 GGAGACACCATGTTGACAGAAATTTGTACCTGGGGGCCATAGACATGATGATCTGCGA 784
QY      609 LysArgGlyHisLeuValSerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyr 628
Db      785 AAACGTGGCCACCTGGTGGCAGCAGCTGGAGTGGAGTGGTGGTGGTGGTGGTGGTGGTGG 844
QY      629 AlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysVal 648
Db      845 GCCCTCAACTATCTGGAGGACAAAGGCTGCCCCATGTCATGCTCTGCCCGGAGGTG 904
QY      649 LeuLeuAlaArgGluGlyArgAspGlyAsnProProPheLeuLysLeuSerAspProGly 668

```

```

905 CTCCTGGCTCGGAGGGGCTGATGGAGCCCGCCCTTCATCAAGCTGATGACCCCTGGG 964
669 ValSerProThrValLeuSerLeuGluMetLeuThrAspArgIleProTrpValAlaPro 688
965 GTCAGCCCGCGCTGTGTAGCCCTGGAGATGCTCACCGACAGAGTCCCTGGGTGGCCGCC 1024
689 GluCysLeuGlnGluAlaGlnThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAla 708
1025 GAGTGTCTCGGAGCGCGACAGACCTTAGCTTGGAAAGCTGACAAAGTGGGCTTCGGCGCC 1084
709 ThrThrTrpGluValPheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaLys 728
1085 ACGGTCTGGGAAGTGTAGTGGCGTCACCATGCCCATCAGTCCCTGGATCTGCTAAG 1144
729 LysLeuLysPheTyrGluAspGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAla 748
1145 AAACCTCCAAATTTATAGGACCGCGCAGCAGCTGCGCGCCCAAGTGGACAGAGCTGGCC 1204
749 GlyLeuIleThrGlnCysMetAlaTyrAspProGlyArgArgProSerPheArgAlaIle 768
1205 CTGCTGATTCACAGTGCATGGCTATGAGCCGGTCCAGAGCCCTCTCTCCGAGCCGTC 1264
769 LeuArgAspLeuAsnGlyLeuIleThrSerAspTyrGluLeuLeuSerAspProThrPro 788
1265 ATTCTGACCTCAATAGCTCATCTCTTCAGACTATGAGCTCTCTCTCAGACCCACACCT 1324
789 GlyIleProSerProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGln 808
1325 GGTGCCCTGGCACCTCGTGTGGCTGTGG---AATGGTGCCAGCTCTATGCTGCGCAA 1381
809 AspProAlaIlePheGluGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsn 828
1382 GACCCCAAGATCTTCGAGGAGAGACCTCAAGTACATCTCACAGCTGGGCAAGGGCAAC 1441
829 PheGlySerValGluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuVal 848
1442 TTTGGCAGCTGGAGCTGTGCGCTATGACCCGCTAGCCCAATACAGGTGCCCTGGT 1501
849 AlaValLysGlnLeuGlnHisSerValProAspGlnArgAspPheGlnArgGluIle 868
1502 GCCGTGAACAGCTGCAGCAGCGCGGCAGACCAGCAGAGGGACTTTCAGCGGGAGATT 1561
869 GlnIleLeuLysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGly 888
1562 CAGATCTCTCAAGCACTGCACAGTATTTCATTGTCAAGTATCTGCTGCTCAGTATGCG 1621
889 ProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAsp 908
1622 CCGGCGCGCCAGAGCTGGGCTGTCTATGGAGTACCTGCGCAGCGGCTGCTTGGCGCAC 1681
909 LeuLeuGlnArgHisArgGly---LeuHisThrAspArgLeuLeuLeuPheAlaTrpGln 927
1682 TTCCTGACGCGCACCGCGCGCGCTCGATGCGCCGCTCTCTCTATTCCTCGCGAG 1741
928 IleCysLysGlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaAla 947
1742 ATCTCAAGGGCATGAGTACTTGGCTGCGCTGCGCGCTGCGTGCACCGGACCTGGCCGC 1801
948 ArgAsnIleLeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLys 967
1802 CGAACATCTCTGTGAGAGCGGACGACAGCTCAAGATCGCTGACTTCGGCGCTAGCTAAG 1861
968 LeuLeuProLeuGlyLysAspTyrTrpValValArgGluProGlyGlnSerProIlePhe 987
1862 CTGCTGCGCGCTTGACAAAGACTACTAGCTGGTCCGCGAGCCAGGCGAGCCCAATTTTC 1921
988 TrpTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSer 1007
1922 TGGTATGCCCCGAAATCCCTCTCGGACACATCTCTCTCGCAGTCAGACCTGCGAGC 1981
1008 PheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAla 1027
1982 TTCGGGGTCTCTGTACGAGCTCTTCACTACTCTGCGACAAAAGCTGCGAGCCCTCGGCC 2041
1028 GiuPheLeuArgMetMetGlyProGluArgGluGlyProProLeuCysArgLeuLeuGlu 1047
2042 GAGTTCCTCGGATGATGGATGTGAGCGGATGTCCCGCCCTCTCGCCCTCTCTGGAA 2101
1048 LeuLeuAlaGluGlyArgArgLeuProProProProThrCysProThrGluValGlnGlu 1067
2102 CTGCTGAGAGCGGCGCAGAGCTGCGCGGCTCTCTCGCTGCGCTGCTGAGGTTCAGGAG 2161
1068 LeuMetGlnLeuCysTrpAlaProGluProHisAspArgProAlaPheAlaThrLeuSer 1087
2162 CTATGAAGCTGTGCGGCCCTTACCCACAGGACCGGCCCATCTTCAGCGCCCTGGGC 2221
1088 ProGlnLeuAspProLeuTrpArgGlyArgProGly 1099
2222 CCCAGCTGACATGCTGTGGAGCGGAAGCGCGGG 2257
HSU70065 13562 bp DNA linear PRI 27-JAN-1997
Human JAK3 gene, complete cds.
U70065
U70065.1 GI:1800224
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 13562)
Riedy, M.C., Dutra, A.S., Blake, T.B., Modi, W., Lal, B.K., Davis, J.,
Bosse, A., O'Shea, J.J. and Johnston, J.A.
Genomic sequence, organization, and chromosomal localization of
human JAK3
Genomics 37 (1), 57-61 (1996)
97079659
8921370
2 (bases 1 to 13562)
Riedy, M.C., Dutra, A.S., Blake, T.B., Modi, W., Lal, B.K., Davis, J.,
Bosse, A., O'Shea, J.J. and Johnston, J.A.
Direct Submission
Submitted (09-SEP-1996) NIAHS, National Institutes of Health, 9000
Rockville Pike, Bethesda, MD 20892, USA
Location/Qualifiers
1. 13562
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.1"
201. 13562
/gene="JAK3"
join(201. 384,719. 842,1141. 1255,1462. 1607,2025. 2317,
2865. 2987,3082. 3350,4264. 4450,6303. 6430,6630. 6761,
7418. 7503,7918. 8044,8808. 8939,9019. 9171,9302. 9452,
10839. 10978,11059. 11256,12009. 12125,12400. 12969)
/gene="JAK3"
/codon_start=1
/product="JAK3"
/protein_id="AAC50950.1"
/db_xref="GI:1800225"
/translation="MAPSEETPIIPQRSCLLSTEAGALHVLIPARGPPQRLSFS
SGDHLAEDLCQAAKAGSILFVHSLFALATEDLSWFPFPHSHIISVEDASTQVLLFRI
RSFVFNWFLGKCHRFGLRDLASAILDLFVHLFAQHSRDLVSGSLVGLSUKQ
GECILSLAVLDLARMAREQAQPGELLKTVSKACLPSPRLDIQGLSFVTRRAIRTV
PRALPRVAACADBSHLSQAKYIMDLERDPAGAAETPHVLPALGGHGLGLFRVAG
DGGIANTGEGVLQPPCFPRIVDISIKQAPRGVPGHRLVTVTRTDPALEAKLTG
LPEALSVFALVDGFRLLTDSQHFCEKVEAPRLLLEVAEQCHGPTIDPALEAKLTG
GSRGSLYVLRSPQDFLLTVQVQNPGLKLRSPETGKLRSPETGKLRSPETGKLRSP
RELLATGWDGLHVDVAVLTSCCIPRKEKSNLIVVQRGSPPTSSLVQSQQL
SCMTPHKIPADSLHWHENLHGHSFTKIYRGCRHEVVDGEARKEVLLKMDAKHNCM
ESFLFAASLMSQVSRHLVLLHGVCMAGDSTMVQFVHLGALDMLYLRKRGHLVPSWK
LOVQKQALVALNYLEDKGLPGNVSARKVLLAREGADGSPFIIKLSDPGVSPLVSL
MLTDKIPWPAECLREAOITLSLEADKNGFGATVWVFGVMPISALDPAKQLQFVED
FEATURES
source
gene
CDS

```


Db 2599 AGAGCTGGAGGCTGCGAAGCGGGGCGAGGATCAAGGGCGGGTCCGGCAGAGAGAAGG 2658
QY 278 -----TtpSerSerAsnAspGluLeuPhe----- 285
Db 2659 AGCGGGGATTTGCAGAGCGGTGGGAGATGTTTGGGATGGGTAGGAGGTGTGAGA 2718
QY 285 ----- 285
Db 2719 GGTGGAGTTGATAGGAGGGGCGCGCTTGGAAAGGTTGAATGGCAAGGATAGGAG 2778
QY 285 ----- 285
Db 2779 TGGATGGTGTGCTTGGGGTGGTTTCATGGCGTGGTTTGGCGGACCCGCTGGGCC 2838
QY 286 -----GlnThrPheCysAspPheProGluLeuVa 295
Db 2839 CACTACGGTAANCCCCAACCAACCACTCTCCAGCCCTCTCGGACTTTCCAGAAATCGT 2898
QY 295 lAspValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValTh 315
Db 2899 AGACATTAGCATCAAGCAGGCCCGCGGTGGCCCGCGGAGAGACACCGCTGTGTAC 2958
QY 315 rValThrArgMetAspGlyHisIleLeu----- 324
Db 2959 TGTACCAGGACAGACAACCAAGATTTAGGTGTGTGAGGATTCCTCCCTCTCAGCCT 3018
QY 324 ----- 324
Db 3019 TACCCGAGGGCGGACCGGACCCCTCGGTTTCACTGGGCTCTGACGCTTGTCCCTCG 3078
QY 325 --GluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyT 344
Db 3079 CAGAGCCGAGTTCCAGGCGTCCCGAGGCTCTGTCTGTGGCGCTCTGGACGGCT 3138
QY 344 YrPheArgLeuIleCysAspSerArgHisTyPheCysLysGluValAlaProProArgL 364
Db 3139 ACTTCGGGTGACCAAGGACTCCAGCACTTCTTGCAAGAGGTGGCACCGCGGAGGC 3198
QY 364 euLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisL 384
Db 3199 TGCTGGAGGAGTGGCGGAGAGTGCACGCGCCCATCACTCTGGACTTTGGCATCAACA 3258
QY 384 ysLeuLysAlaAlaGlySerLeuProGlyThrTyIleLeuArgArgSerProGlnAspT 404
Db 3259 AGCTCAAGACTGGGGCTCACGTCCTCGCTCTATGTTCTCCGCGCAGACCCCGAGACT 3318
QY 404 YrAspSerPheLeuLeuThrAlaCysValGln-Thr----- 415
Db 3319 TTGACAGCTTCCTCCTCACTGTCTGTGTCCAGGTGGTCTACTGTAGGGTGGGTAGTGG 3378
QY 415 ----- 415
Db 3379 AGGGCTGCCTGGACAAGAGTTACCTCTGAATTGAGATTTACTCAGATCTACTCAGCAT 3438
QY 415 ----- 415
Db 3439 TTGGTTCTCGAACAATAAGGGAACGACACCCCGGTTAAGCGACACGCTATTCT 3498
QY 415 ----- 415
Db 3499 CTGGTCTAAATACTGGAATTAGTTGGTGTGCTTCAGGATAAAGACTGAAGCTGGGCGC 3558
QY 415 ----- 415
Db 3559 TGTCCGGTCCCTGTTATCCAGCCCTTTGGAGGCTGACGTAAGCATATCCCGAGATC 3618
QY 415 ----- 415
Db 3619 GGGATATCCACCCATCTGGTCCAAATGGTTAAACCCCTGTCCACCTGTAATAATGTAA 3678
QY 415 ----- 415

Db 3679 CTGAGCTGGGCATGGTAATGGGCATCTGTTAATCCACCTACTCGGGAAGCTCAAGGCAT 3738
QY 415 ----- 415
Db 3739 GTTTTCTAAGGTTTTTAAAAAAGGGGGGGGTTTTTTTGGGGTTTTTATGGGAAAAAAA 3798
QY 415 ----- 415
Db 3799 AAGGGGTGGGGNAAAAAAGGGTAACCCCTTTTTTTAGTCTTCGAGAGCTCTCTGGCAAC 3858
QY 415 ----- 415
Db 3859 CTGCTGGGATGGGGGCTGCACGTAGATGGGTGGCAGTGACCTCCTCCTCTCTGTGTAC 3918
QY 415 ----- 415
Db 3919 CCCAAATTTGGGGGTGGGTGGCCAGTTTAAAGGTTTAAAGGTTTGNAACTTCGGG 3978
QY 415 ----- 415
Db 3979 GCCCAAAGGGGAACCCCTTTTTTTTTTAAAAAATTTAAAAAATTTTGGGGGGGTAAAG 4038
QY 415 ----- 415
Db 4039 GGGATTTTTTTTCCCAAAATTTGGGGGGGGCGGGGCATTTTTTTAAACCCGGGG 4098
QY 415 ----- 415
Db 4099 GGGGGGGTTCCAGGGGCCCAGGATAAACCTTTAAATTTCCCGGGGAAAAAGGCCCA 4158
QY 415 ----- 415
Db 4159 AATTCTTTTCCCAAAAAAAGAAAAAAGAAAGAAAGAAAGAAAGAAAGAGAG 4218
QY 416 -----ProLeuGlyPro 419
Db 4219 AGTGTCTGTGCTCAGTCCCACTCAGGGGCCACTCTTCTTTGCAGAACCCCTTGTCT 4278
QY 420 AspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeu 439
Db 4279 GATTATAAGGCTGCTCATCGGGCGAGCCGCCACAGGAACCTTCTCTGTGTGGCCCTC 4338
QY 440 SerGlnProHisArgSerLeuArgGluLeuAlaCysTyrAsnSerGlyLeuArg 459
Db 4339 AGCCAGCCCAACAGCAGTCTTCGAGAGCTCTCGGCAACCTCTGTGGATGGGGGTGAC 4398
QY 460 ValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArg-ProLys----- 476
Db 4399 GTAGATGGGTGGCAGTGACCTCCTCCTCTGTATCCCGCAGACCCCAAGGTGAGGCC 4458
QY 476 ----- 476
Db 4459 CCTGGTCCCTGGAATGAGTGGCTGATCTGGGACCTGGTTTCTCTATGTCTGATAACAG 4518
QY 476 ----- 476
Db 4519 CTCCTGTGTGATGGCAAGTGTAGAAACTGCAGATCATGTGTGGTTAGGGAAGAAAGG 4578
QY 476 ----- 476
Db 4579 CGATTTGTGCTTCAGGAAGTTAGAGCCATCTAACCCCTAGGGCTGGATTGCCACGGTA 4638
QY 476 ----- 476
Db 4639 CAGCTAGATGTAGCCATGTACCTGGACCTTGNNTTATCTCCAGAACTCTCAGTCTTTC 4698
QY 476 ----- 476
Db 4699 TTCTGGGTGACCTTCTCCTCTTGGCACAACCTCCCGCTGTATGGGGAACAAATTAATT 4758
QY 476 ----- 476
Db 4759 GAGTATGTTGAGCCCTTTCTCTGAGTTCTTCCAGCAATTCCTCGGTGAAGAAACCCGAT 4818

```
QY 476 ----- 476
Db 5899 GTGGTCTCTCGATAAAACACAGAAATTTCCCGGGGAAACCCGGGAATTTTTT 5958
QY 476 ----- 476
Db 5959 TTTTAAAGAACAAAAAATAGCCGGGACCGGTGGCACACACCTGTAGTCCAGC 6018
QY 476 ----- 476
Db 6019 TATTCGGAGCCGAGGACAGAAATCGCTTGAACCTGGGAGGTGGAGTTGCAGTGAGC 6078
QY 476 ----- 476
Db 6079 TGAGATCGACCATTTGCCACCCAGCCTTGATGACAGAGTGTAGATCCGTCTCAGCAGCAG 6138
QY 476 ----- 476
Db 6139 CAGCAACAGCAAAACAAAAACAACAAAAAGCATGTGCTGAGTCTTATCTCAG 6198
QY 476 ----- 476
Db 6199 GGTGGGTCTAGAGGTACCTCAAACTAAGCATGAGTTAGCTAACCTTGGGACTTT 6258
QY 477 ----- GluLysSerAsnLeu11 482
Db 6259 TCACCTCTGATTTCTGGTCTTCTCCCTCATCTCTCCCATAGAAAAGTCCAACTGT 6318
QY 482 eValValArgGlyCysAsnProAlaProAlaProGlyCysSerPro---SerCysCy 501
Db 6319 CGTGGTCCAGAGAGGTACAGCCACCCACATCATCTCTGGTTCAGCCCAATCCCAATA 6378
QY 501 sAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrp 518
Db 6379 CCAGCTGAGTCAGATGACATTTCAACAAGATCCCTGCTGACAGCCTGGAGTGGGTAAGAG 6438
QY 518 ----- 518
Db 6439 CCCTGGGAAATAGCGCATACCTCAGTCTGGGGTCCAGAGACTCAGATCGGTGGCTCAG 6498
QY 518 ----- 518
Db 6499 GCATATGCTATAATTTTACCTTGGCTCGGTTCCTGATCTGTAAATGGGCCAGCAGCT 6558
QY 518 ----- 518
Db 6559 ATGTCTCGCTGGGCTGGATCCTGCAGGAACCCCTCACTGGCTCTTCTGTCTCCCC 6618
QY 519 ----- HisGluAsnLeuGlyHisGlySerPheThrLysIlePheHargGlySerA 535
Db 6619 TCACCATTCAGCATGAGAACCTGGGCCATGGGTCTTCCACCAAGATTTACCGGGGCTGTC 6678
QY 535 rGArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAsps 555
Db 6679 GCCATGAGGTGGTGGATGGGAGGCCCGAAGACAGAGAGTCTGCTGAGGTCTATGGATG 6738
QY 555 erArgHisArgAsnCysMetGluSer 563
Db 6739 CCAAGCACAGAACTGCATGGAGGT-GAGAGGCAATGTGGCAGACTTTTGGAGTCCGG 6797
QY 563 ----- 563
Db 6798 GGCTGGCTTGAGCGGGGCTCGTGGATGCAGAGACATTTACAACACACAGGGAACCTGGC 6857
QY 563 ----- 563
Db 6858 GTGGTGGCTGATGCTGCATCATAGCGGGACTCTGCTTGCACAAAAAATTTTAGGCC 6917
QY 563 ----- 563
Db 6918 GGGCGCGGTGGCTCAGCGCTGTATATCCAGCAGCTTTGGAGAGCTGGAAGCGGCAAGATCA 6977
QY 563 ----- 563
```


Db 6978 TCAGGTGGGGAGTCTCGAGACCAACCTGCCCAACATGAGAAACCCGCTCTCTACTAA 7037
QY 563 ----- 563
Db 7038 AAAACCCAAATTAACCGGGCGTGGTGTGATGCTGTAAATCCAGCTACTCGGAGTC 7097
QY 563 ----- 563
Db 7098 TGAGGCAGGAGAAATTCCTTGAATCTGGAAAGCGGAGTTCACAGTCAAGATCGGCC 7157
QY 563 ----- 563
Db 7158 ACTGCATCCAGCTGGTGACAGAGGAGACTCTGTCTCAAAAACAAANCAANCAAAA 7217
QY 563 ----- 563
Db 7218 CAAAAACCATAGATGATAGTGGAACTTCTGTCCCGTATCAGAAAAATCATGTAGTGCT 7277
QY 563 ----- 563
Db 7278 GTGTGCATTAATGGCAGACTCCAGGGCCAAAGGTGACCTGTGCCAGGTGTCCCTTAAG 7337
QY 563 ----- 563
Db 7338 GCAGTCTGTGAGCAAAAAATTTGGATTATTGGAGTGAAGAAACCCAGGCATCTTCTC 7397
QY 564 ----- PheLeuGluAlaAlaSerLeuMetSerGlnValSerT 576
Db 7398 TCCCTTCCACCTTCCCGAGTCATCTCTGGAAGCAGCGAGCTTGATGAGCCAAAGTCTGT 7457
QY 576 YrProHisLeuValLeuHisGlyValCysMetAlaGlyAsp----- 590
Db 7458 ACCGGCATCTGTGTCTCTCCAGCGGTGTGCATGCTGGAGACAGTGCTGCTCCTAG 7517
QY 590 ----- 590
Db 7518 GGGGTGTTAAAGGGAGAGAGGAGAAAGAGGCGAGGGGGAAGAGAGTAGGGGCA 7577
QY 590 ----- 590
Db 7578 GAGAGAGAGAAAGCCGGGAGAGGCCAACCGGAGCAACCAAGAGAGGTTAGCGAGA 7637
QY 590 ----- 590
Db 7638 GGATAAAAACTGGCTGACGTTTATAAAGATAGGGGGAGAGCCCTGCCTCGAGC 7697
QY 590 ----- 590
Db 7698 AAATTATATAGCCCGGAGAGAGCTTCTGGCTTGGGGGGGCTGAGAGGAGACCC 7757
QY 590 ----- 590
Db 7758 GCCCCTTTAACAAACTAGAAAGACCTCTGGGGATAATCCCGCTTATTATAAACCTGG 7817
QY 590 ----- 590
Db 7818 AGAGAACTCAACGAAAAAGAGATCCCTGGGATGCCCCAGATATATTTGGGAGTGCC 7877
QY 591 ----- SerIleMetValGlnGluPheV 598
Db 7878 AAAGAACTAAACTTTGGAATTAAGCAAAATATGGCTGAGCACCATTGCTGAGGAATTTG 7937
QY 598 alTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerT 618
Db 7938 TACACTGGGGCCCATAGCATGTATCTGCGAAAAAGTGGCCACCTGTGTGCCAGCAGCT 7997
QY 618 rpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyr----- 632
Db 7998 GGAAGCTCAGGTGCTCAACAGCTGGCTAGCCCTCAACTATCTGTGTAGTGCTCCTC 8057
QY 632 ----- 632

Db 8058 CCCCCAGGGATTGTGTTGGGGGAGGGGGGTTTAAAAAACCCAAACCCCTAAAT 8117
QY 632 ----- 632
Db 8118 TTTGGGGGGGGTGTGTGGGCTGTTCCTCCCTCCATTCCCGGAGAAAGTTTTTTTTTTTT 8177
QY 632 ----- 632
Db 8178 AAAGGGGAAAAGGAATGTTTTTTTTTTTAAAGATTGTGTAGGTTTGGGGGGGG 8237
QY 632 ----- 632
Db 8238 CATTTTAGTTTCTTTTTTGGGAAAATGGNAGAAAAATTTTTTTTTTTTTTTTNGGTT 8297
QY 632 ----- 632
Db 8298 TTTGGGAAAAGGGGTGGCCGGGAGAGAGGTTTCTTTTTTGGAGAAAAGGACATGCC 8357
QY 632 ----- 632
Db 8358 CTTTCACTGGAGTGGGTTTTTTCATGAAAAAGAGGAGGATTTCCTCGGCCCTATCAGCATT 8417
QY 632 ----- 632
Db 8418 ACCAAAAGCAAGCCCGCGGAGAGAAAGGAGATATTTCCTCGGCCCTATCAGCATT 8477
QY 632 ----- 632
Db 8478 TCCGTTGCCAGGAGACAGAGGAAAAGTGGGAAAGGAATTAATTGGAAGCTTGGAGCAG 8537
QY 632 ----- 632
Db 8538 GGCCCTGGAATCCAGGTGAGGAGGTTAAATTTTTTTTCAAAGGCATTGGGGTTAACAG 8597
QY 632 ----- 632
Db 8598 ATGGGTGTGTGTTTGAGAGGGGAGGACATACATGTCATTTTTGGGGGTGAAGTAGAG 8657
QY 632 ----- 632
Db 8658 GGGCAAGACAGGATATCAGATTTCAGGACGATGTGTGGAGCATGTTTGAGCAGTACCAA 8717
QY 632 ----- 632
Db 8718 GTGGGTTTGAAGGATATAGGAGTTTCCAAACAGACATCTTATTTATCAACCTCC 8777
QY 633 ----- Leu-GluAspLysGlyLeuProHisGlyAsnVal 643
Db 8778 CGGGCATTTTCGTGTGTTTGGCCCCCTTAGAGGACAAAGGCTGCCCATGGCAATGTC 8837
QY 644 SerAlaArgLysValLeuLeuAlaArgGlyGlyAspGlyAsnProPheIleLys 663
Db 8838 TCTGCCGAAAGTCTCTCTGCTCGGAGGGGCTGATGGAGCCGCCCTTCATCAAG 8897
QY 664 LeuSerAspProGlyValSerProThrValLeuSerLeuGlu----- 677
Db 8898 CTGAGTGACCTGGGTCAGCCCCGCTGTGTAAAGCTGGAGTAAGTTCTCTGGAGGTGG 8957
QY 677 ----- 677
Db 8958 AGAGGGGGGCTCAACAGAGGCAAGGAGTGATCCCTGATCCCACTTTCATTCCTCCCA 9017
QY 678 -MetLeuThrAspArgIleProTyrValAlaProGluCysLeuGlnGluAlaGlnThrLe 697
Db 9018 GATGCTACCCAGACAGGATCCCTGGTGGCCCCCGAGTGTCTCCGGAGGGGCGACACT 9077
QY 697 uCysLeuGluAlaAspLysTyrPheGlyAlaThrThrTrpGluValPheGlnArgL 717
Db 9078 TAGCTTGAAGCTCAAGTGGGCTTCGGGCCACCGTCTGGGAAAGTGTATTAGTGGCGT 9137
QY 717 yProAlaHisIleThrSerLeuGluProAlaLys----- 728
Db 9138 CACCATGCCCATAGTGGCTGGATCTGTCTAGGTCAGAGCCCTCACCCGGCATCGGT 9197

QY	728	-----	728	QY	778	-----	778
Db	9198	CTCCGAACCCCACTTTGACAGCCGGGACAGTGCATCCAGTCTGGGGAGATTGGGGTG	9257	Db	10278	TCCACACAAAGAACCCAAAAAGACCTAGCACATCATTTCAAAGTGTGAATGTGCAATT	10337
QY	729	-----	734	QY	778	-----	778
Db	9258	GGTCTATTGGTTGGGATTACCGACTGCTCCTCTCACCTCAGAACTCCAATTTATG	9317	Db	10338	CTTTCTCAATTTTGGGCTGTCCCTCTTCAAAATTCATCTTTTCAAAGGGCAATTTG	10397
QY	734	luAspGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnC	754	QY	778	-----	778
Db	9318	AGACCGGACGACGCTGCCGCCCCCAAGTGGACAGAGCTGGCCCTGCTGATTCAACAGT	9377	Db	10398	AAGAGTATTACGGTAGAGGAGGTGTCTCTATCTCGTCCAACTATGGGCAATGGCA	10457
QY	754	ysMetAlaTyAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnG	774	QY	778	-----	778
Db	9378	GCATGCCATATAGCCGGTCCAGAGGCCCTCTCCGAGCCGTCATTCGTGACCTCAATA	9437	Db	10458	GAGTGACCGTGAATATATAGGCAATTTTGATCGAGTCGACGCCCTATAGTCGATAG	10517
QY	774	lyLeulleThrSer-----	778	QY	778	-----	778
Db	9438	GCCTCATCTCTTTCAGGTGCCCGCTGGGACGGGTTGGGTGGGGGGCTGTGATGTCATAT	9497	Db	10518	AGCTTAGTATATAGAACCGCAGGAAATTTGATGTCTTCAATTTATGCCCCCCCCATCC	10577
QY	778	-----	778	QY	778	-----	778
Db	9498	TGGGCCAGTGAAGGAGCGTGTTCAGTAGGACACGCGCTGTGTCTGTGTGAGGTT	9557	Db	10578	TTTTGTAAACCAATTTGGGAAGTTGGCACCAGGTTTTCACCTGCACGAGGAGGCC	10637
QY	778	-----	778	QY	778	-----	778
Db	9558	GGAGGGGTTGATGACTATGACAGACAGTGTGAGCTTCAACAGTACTGTAAACGAACAGTC	9617	Db	10638	TTCTTTTCCCTGGAAACAAGAGTGGGAAGACACCTTTTTTTTCCAGTCTCTCCACAAGGT	10697
QY	778	-----	778	QY	778	-----	778
Db	9618	CCCTCAGTTCAGGTTCAACATGATGATATTTATTTATTTATTTATTTTGGAGACAGAG	9677	Db	10698	ACCTTTTGGCTTGCAATGGCCCTGCCATTAATGCACAGAGGGTCAATATGCAGATGGA	10757
QY	778	-----	778	QY	778	-----	778
Db	9678	TCTTGCTCTGTGCTAGGCTGGAGTGCAGTGGGACAAATTCGGCTCAGTGGAACTCTA	9737	Db	10758	GGTTGCACACCAAGTCAAACTCAGAGTGGGACCCAGGATGAGAGGGGCTGTTACCACT	10817
QY	778	-----	778	QY	779	-----	792
Db	9738	CCTCCTTGAACCTCTAGTTCAAGCGATTCTCCTGCTCAGACTCCTGAGTAGCTGGGATT	9797	Db	10818	GCCCATGCCCCACCCCAAGTATGAGTCTCTCAGACCCCACTGGTGGCTGGT	10877
QY	778	-----	778	QY	792	erProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyAlaCysGlnAspProAlaI	812
Db	9798	ACAGTGTGTGCCACCACGCCAGCTAATTTTGGTTTGGGGAAGATTTTGGGGT	9857	Db	10878	CACCTGTGATGTCTGTGG---AATGGTCCCCAGCTCTATGCTGCTCCAGACCCCAAG	10934
QY	778	-----	778	QY	812	lePheGluAlaArgHisLeuLysTyrlleSerLeuLeuLys-----	826
Db	9858	TTCAACCATATTTGGCTAGGGTGTCTTCGAACCTTTCACCCCAAGGTGATCCGCCCCAC	9917	Db	10935	TCTTCGAGGAGAGACACCTCAAGTACATCTCACAGCTGGGCAA-GGTAAGGTGGCAGGG	10993
QY	778	-----	778	QY	826	-----	826
Db	9918	CTCTACCTCCTAAAGTGTGGGATTCAAAATTTCTGCTACTTGGCTTTGTTTGTATAT	9977	Db	10994	CCAGGTGGTGGAGAGGACGAGCATCCAGGTACCTGGACATCAGTCCCGCTATCC	11053
QY	778	-----	778	QY	827	-----	845
Db	9978	TGCCAAATTTAGTCGAGATATAGTTTGGCAGAGTAGAGTCATGATTTTCAGTTCGTGAG	10037	Db	11054	TCGAGGGCACTTTGGCAGCGTGGAGCTGCGCTATGACCCGCTAGCGCAATACAG	11113
QY	778	-----	778	QY	845	lyProLeuValAlaVallylsGlnLeuGlnHisSerValProAspGlnGlnArgAspPheG	865
Db	10038	ACTTTGTAAACAAGAGTTTCAAGTGTCAAAGGAGGAAACCGGATTGTTTATAATAAAA	10097	Db	11114	GTGCCCTGTGGCCGTGAAACAGCTGCAGCAGCGGGCCAGACCAGCAGAGGGACTTTC	11173
QY	778	-----	778	QY	865	lnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVallylsTyArgGlyV	885
Db	10098	CCTGCAGTATTAAATAGGAATCATAATAATGACTCAGTATAATACCAGGAGGAAACAC	10157	Db	11174	AGCGGAGATTTCAGATCCTCAAAGCACTGCACAGTATTTTATTGTCAGTATCGGTG	11233
QY	778	-----	778	QY	885	alSerTyGly-----	888
Db	10158	CCAAATAAGTTTAAAGTTGCCCTAAAGAAAGAAATGGGTGATGGTAAATAATTTA	10217	Db	11234	TCAGCTATGGCCCGGGTGAGCCAGCTCCCGGATGAGTGAACCAAGACGTATGGTCTTT	11293
QY	778	-----	778	QY	888	-----	888
Db	10218	AAGTCCCGCAACAAATTAAGCGTATAGATTCTGTAAACTGAAACATAAGCAAAA	10277	Db	11294	TCAAAGTGCACATTTCTTACCCTCTGCGAGGCCACTTTAGTAGGTGGAAACGTGATTA	11353
QY	778	-----	778	QY	888	-----	888


```

OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 7..3906
FT FT /*tag= a
FT FT /product= JAK3 protein-tyrosine-kinase
FT FT /note= "EC-2.7.1.112"
XX
XX
PN W09601838-A1.
XX
XX
XX 25-JAN-1996.
XX
XX 28-JUN-1995; 95WO-US08354.
XX
XX 08-JUL-1994; 94US-0272368.
XX
XX (UTEM ) UNIV TEMPLE.
XX
XX Rane SG, Reddy PE;
XX
XX WPI; 1996-097584/10.
XX P-PSDB; AAR88560.
XX
XX JAK3 protein tyrosine kinase and DNA encoding it - useful to induce
XX differentiation of haematopoietic cells to treat or control
XX leukaemia
XX
XX Claim 4; Page 26-28; 49pp; English.
XX
XX The sequence encodes a novel protein-tyrosine-kinase of the Janus
XX kinase (JAK) family (involved in cytokine signal transduction),
XX designated JAK3. The sequence has been isolated from mouse 32Dcl3
XX cells growing in interleukin-3 or granulocyte-macrophage colony
XX stimulating factor by reverse transcription-polymerase chain
XX reaction amplification, using primers J3 and J4 (AAR11084-111085),
XX based on conserved sequences in JAK JH2 and JH1 domains. A 950-bp
XX cDNA probe is isolated, which is then used to screen a granulocyte
XX colony stimulating factor-treated 32Dcl3 cDNA library in phage
XX lambda-gt11 in Escherichia coli to isolate the full-length gene.
XX The JAK3 protein is useful for stimulating the differentiation of
XX haematopoietic cells in therapy or control of leukaemia.
XX
XX SQ Sequence 4016 BP; 742 A; 1203 C; 1235 G; 836 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 4016
Score: 5223.00 Matches: 1060
Percent Similarity: 80.59% Conservative: 3
Best Local Similarity: 80.36% Mismatches: 32
Query Match: 89.13% Indels: 235
DB: 17 Gaps: 6

US-09-397-967A-16 (1-1099) x AAT11083 (1-4016)
QY 1 MetAlaProSerGluGluThrProLeuileProGlnArgSerCysSerLeuSerSer 20
Db 7 ATGGCACTCCAGTGGAGGACACCTCTGATCCCTCAGCGCTCTTGACGCTCTCATCC 66
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
Db 67 TCAGAGGCAGAGCCCTGCAATGTCCTCTCTCCCGGGGACCTGGGCGCTCCGACGGA 126
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaIysAla 60
Db 127 TTGTCAATCTCTTTTGGGGACTACTTGGCTGAGGATTTATGTGGGAGCTGCCAAGGCC 186
QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db 187 TGTGGATCCCTGCTGTTTATCATTCCTGCTTTTGGCTGCGGCACTGAGGACTTCTTTGC 246
QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db 247 TGGTTTCCCCCAAGCCACATCTTCTGCATAGAGGACGTGGACATCTCAAGTCTTGGTCTAC 306

```

```

QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCystHisArgPheGlyLeu 120
Db 307 AGGCTACGCTTTATTTCCCTGACTGGTTTGGCTGGAGACATGTCCAGGCTTTGGGCTG 366
QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
Db 367 CGCAAGATTTCACCAAGTCCCATCTTGTGACTTACATGTTTATAGAAATCTCTTTGTCT 426
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
Db 427 CACCGCAGTGACCTGTGTGAGTGGGCGCTCCCGTGGGCTTAGCATGAAGGACAGGGA 486
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
Db 487 GAGTTCTTGAGCTGCGCTGCTGGACTTGGCCAGATGGCTCGTGAGCAGGCCAGCGC 546
QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
Db 547 CCAGGAGAGCTGCTGAAGACGGTCAGTTACAAAGCCTGTCTGCCGCCAGCCTGCCGAT 606
QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAla 220
Db 607 GTGATCCAGGCCAGAACTTCTGTCACAGCAGCGCATCCGCAAGGCGCTGTCTTGGCG 666
QY 221 LeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeu 240
Db 667 CTGC--GCGTGTGTCGCTGCCAGCCGCTACGGCTCATGGCCAAAGTATAT-CTG 723
QY 241 AspLeuGluArgLeuHisProAlaAlaThrGluThrPheArgValGlyLeuProGly 260
Db 724 GACCTGGAGCGGCTACATCCAGCGGCCACACCGAGACCTTCCGTGTGGGGCTCCCGGGC 783
QY 261 AlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrSerSer 280
Db 784 GCCCAGGAGGAGCCGGGGCTTTCGCTGTGGCGGGGACAAACGGCATCTCTCTGGAGTCC 843
QY 281 AsnAsp----- 282
Db 844 GGGGACCAGGAGGTGCTGGGCTTGGAGAGCGGGTCCCGCGGGGAGCGCTGGCGG 903
QY 282 ----- 282
Db 904 TTAGTAAGAAAGCCCGTCCGAACAATCAGAAAGTGCTAGTCCCGGGCGGGGTGATTCG 963
QY 282 ----- 282
Db 964 ACTACGGGCGAGGGGGCGGGGTCAATTAAAGTGGGTCCGGGATGGAGCTGCCCCAG 1023
QY 282 ----- 282
Db 1024 GGACTGACCTGGGCGGTGACTAGAAAGGTGAGACTAGATAGGGGGCGGGACGACAGAA 1083
QY 282 ----- 282
Db 1084 GGCATAGTATGGATTGGATCAGTGGCATGACCCGACAGGCGCTGTCTTCAGTCCCCCTA 1143
QY 282 ----- 282
Db 1144 ACCTGTCCTCCCCCCCCCACAAGTGGAGGTGGGAAGAGGAGGAGGCTGTGCT 1203
QY 282 ----- 282
Db 1204 GGGTCAAGGTCTGTAATCCCTGGCTTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1263
QY 282 ----- 282
Db 1264 TTTGCGACACAGGTTTCTCGTATAGCTCTGGTGTCTGTGAACTCACTTTGTAAGACA 1323
QY 282 ----- 282
Db 1324 GGCTGGCCCTCGAACTCAGAAATCCGCTGCTCTGCTCTCTGCTCTCTGAGTGTGATTAAAGG 1383

```

QY 283 -----GluLeuPheGlnThrPheCys 289
 Db 1384 CGTGTGCCACTACCGCGCTACTGCCTGCTCTTCTTCTCAGCTTTTCCAGACCTTCTGT 1443
 QY 290 AspPheProGluLeuValAspValSerIleAsnGlnAlaProArgValGlyProAlaGly 309
 Db 1444 GACTTTCCGGAATCGTGAATGTCAGCATCAACAGC- -CCAGTGTGGTCCGGCAGGG 1501
 QY 310 GluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPhePro 329
 Db 1502 AG-CACCGCGTGTCTACTGTACCAAGATGAGCGGCACATCTTGAAGCGGAGTTTCGG 1560
 QY 330 GlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIleCys 349
 Db 1561 GGGCTGCTCAGCGCTGTCTTTCGTGGCCCTCGTGGATGGGTACTTCCGCTGATCTGC 1620
 QY 350 AspSerArgHisTyrPheCysGlyGluValAlaProProArgLeuLeuGluGluAla 369
 Db 1621 GACTCCAGGCATTAATTCTGCAAGAGAGGTGGCGGCCACGGCTGTGGAGGAGGAGCG 1680
 QY 370 AspValCysHisGlyProIleThr----- 377
 Db 1681 GAGCTGTGCCATGGACCATACGTTAGGAGCTAGTTGGGTCAACAGTGGGATGGGGA 1740
 QY 378 -----LeuAspPhe 380
 Db 1741 TGGGCAGCGAGGACTGTCTTGTGGTTGGACCTGGCTGCTGTTTGGCGCTAGGTATAGACTTT 1800
 QY 381 AlaIleHisLeuLeuValAlaGlySerLeuProGlyThrTyrIleLeuArgArgSer 400
 Db 1801 GCCATCCAAAGCTGAAGCGCGCTCGTCTCTCCAGGACCTATATTCTCCGCGCAGC 1860
 QY 401 ProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAsp 420
 Db 1861 CGCAGGACTATGACAGCTTCTCTTACCGCTGGCTCCAGACTCTCTTGGCCCCGAC 1920
 QY 421 TyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSer 440
 Db 1921 TACAAGGGCTGCTCATCCGCCAGGACCCAGCGGGCTTCTCCCTGGTTGC-CTCAGC 1979
 QY 441 GlnProHisArgSerLeuArgGluLeuLeuAlaCysTyrAsnSerGlyLeuArgVal 460
 Db 1980 CAGCCCCACAGAAAGCTCGCGAGGTGCTTGTCAGC-TGCTGGAATTC-GGGCTGGAGTA 2037
 QY 461 AspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGlyLysSerAsn 480
 Db 2038 GACGCTGTGCTGCTGAACCTAACATCTCTGCTGCTGCCCTCCAGACCCCAAGAAAGTCCAAAT 2097
 QY 481 LeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCys 500
 Db 2098 TTGATCGTGTGCGAAGGGCTGACACCCCGCGCTGCGCTGGCTGCTCCCGCTCTGTC 2157
 QY 501 CysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGlu 520
 Db 2158 TGTGGCTGTGACACAGCTGAGCTTCCACACAAATTCACACGGACAGCCTG- -GGACACGAG 2214
 QY 521 AsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValValAsp 540
 Db 2215 AACCTGGTGTACGGTCTTTTACCAGATCTTTCGTGGCCCGCAGCGGGAGTCTGTGAT 2274
 QY 541 GlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCys 560
 Db 2275 GGTGAGACACATGACTCGGAAGTCTCTGGAAGTTCATGGACTCCAGACATCGGACCTGTC 2334
 QY 561 MetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuVal 580
 Db 2335 ATGGAGTCTTTTCTGGAAGCGCAGAGCTTGATGAGCCAAATATCTACCCGACCTGGTG 2394
 QY 581 LeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTyrLeu 600
 Db 2395 TTACTGACGGCGTCTGCTGGTGGACAGCATCATGGTGCAGGAATTTGTGTATCTA 2454
 QY 601 GlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLysLeu 620

Db 2455 GGAGCAATTGACATGATACCTGCCAAGGTGGCCACCTGGTGTACGCCAGCTGAAACTG 2514
 QY 621 GlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHis 640
 Db 2515 CAGGTGACCAAGCAGCTGGCATATGCCCTTAACCTACTTGGAGGACAAAGGCTTCTTCAC 2574
 QY 641 GlyAsnValSerAlaArgLysValLeuLeuAlaAArgGluGlyGlyAspGlyAsnProPro 660
 Db 2575 GGCAACGTCTCAGCACGGAAGGTGCTCTGGCTCGTGAAGGGGTGATGGAAATCCACCT 2634
 QY 661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680
 Db 2635 TTCAITTAAGCTGAGTGTATCTGTGTGTCAGTCCCACTGTGTGAGCTGGAATGCTACCC 2694
 QY 681 AspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 700
 Db 2695 GACAGAAATACCTGGTGGTGGCCCGGAATGTCTCCAGAGGCTCAGACACTCTGCTTGGAG 2754
 QY 701 AlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHis 720
 Db 2755 GCTGACAAAGTGGGGCTTTTGGAGCCACAGCTGGAGGTGTTCAGCGGGGAGCCCGGCCAC 2814
 QY 721 IleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeuPro 740
 Db 2815 ATCACTCTCGTGGAGCCCGCCAAAGCTGAAGTCTATGAGGACCCAGGAGCAGCTGCC 2874
 QY 741 AlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760
 Db 2875 GCTCTCAAATGACAGAACTGGCGGACTTATCACACAGTGCATGGCGTATGATCTGGC 2934
 QY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780
 Db 2935 CGGCGCCCTCTCTCCAGCTATCTCTCAGAGACCTCAACGGCTCATTTACATCAGATTAC 2994
 QY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAla 800
 Db 2995 GAGCTCTCTCCA-GACCCCAACACCTGGCATCCGAGTCCCGAGTCTTCGAGATGAGCTGCGT- - 3049
 QY 801 GlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLysTyr 820
 Db 3050 GCGCGCCAGCTCTATGCTGCTGCCAGGACCCCGCCATATC-GAGGAGAGACACCTTAAGTAC 3108
 QY 821 IleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeu 840
 Db 3109 ATCTCTTTGCTGGGCAAGCGCAACTTTGGCAGCGTGGAGCTGTGCGCTATCACCCCTG 3168
 QY 841 GlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAspGln 860
 Db 3169 ---GACAAATACGGGACCCCTGGTGGCAGTGAACAGCTACAGCACAGCGGGCCAGCCAG 3225
 QY 861 GlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVal 880
 Db 3226 CAGAGGACTTCCAGCGGAGATTTCAGATCTTAAGGCTCTGCACAGGACTTTCATCTGTC 3285
 QY 881 LysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTyr 900
 Db 3286 AAGTACCGGGAGTCAGCTATGGGCCAGGTCCAGAGACCTGCGTTCGTGATGAGTAC 3345
 QY 901 LeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHisThrAsp 919
 Db 3346 CTGCCAGCGCTCGCTCGAGACTTCTCTGAGCGCCATCCGCGGCCCTGCACACCGAC 3405
 QY 920 ArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArgArg 939
 Db 3406 CGCTACTGCTGTTCTGCTTGGCAGATCTGCNAGGCGCATGGAGTACCTGGGTGCGCGCCG 3465
 QY 940 CysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisValLys 959
 Db 3466 TGCCTACACCGCTGACCTGGCTGGCGCAACATCTTGTGGAGAGCGAGGCTCATGTGAG 3525
 QY 960 IleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArg 979

815 Db :||||| 874
260 QY GACCTGGAGCGCTGATCCAGCCGGCGCGGACCTTCCAGCTGGGCTCCCTGG 874
875 Db :||| 277
277 QY :||| 934
935 Db :||| 994
296 QY :||| 316
995 Db :||| 1054
316 QY :||| 336
1055 Db :||| 1114
336 QY :||| 356
1115 Db :||| 1174
356 QY :||| 376
1175 Db :||| 1234
376 QY :||| 1294
1295 Db :||| 1354
416 QY :||| 436
1355 Db :||| 1414
436 QY :||| 456
1415 Db :||| 1474
456 QY :||| 476
1475 Db :||| 1534
476 QY :||| 496
1535 Db :||| 1594
496 QY :||| 515
1595 Db :||| 1654
515 QY :||| 535
1655 Db :||| 1714
535 QY :||| 555
1715 Db :||| 1774
555 QY :||| 575
1775 Db :||| 1834
575 QY :||| 595
1835 Db :||| 1894
595 QY :||| 615

1895 Db GGAATTTGTACACCTGGGGCCATACATATATCTGCGAAACGTGGCCACCTGGTGCC 1954
615 QY rAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAs 635
1955 Db AGCCAGCTGGAAGCTCAGGTGGTCAACAGCTGGCTACGCCCTCAACTATCTGGAGGA 2014
635 QY pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGl 655
2015 Db CAAAGGCTTGGCCCATGGCAATGTCTCTGCCGGAAGGTGCTCTGGCTGGGAGGGGGC 2074
655 QY yAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSe 675
2075 Db TGAATGGAGCCCGCCCTTCATCAAGCTGAGTGACCTTGGGTGAGCCCGCTGTGTAAAG 2134
675 QY rLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGl 695
2135 Db CTTGGAGATGCTCACCGACAGGATCCCTGGGTGGCCCGCCGAGTGTCTCCGGAGGCGCA 2194
695 QY nThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGl 715
2195 Db GACACTAGCTTGGAAAGCTGACAAGTGGGGCTTCGGCGCCACGGTCTGGAAAGTGTTAG 2254
715 QY nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAs 735
2255 Db TGGCGTCAACATGCTGATCAGTGCCCTGATCTCTGCTAAGAACTCCAATTTATGAGGA 2314
735 QY pGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMe 755
2315 Db CCGGAGAGAGCTGCGGGCCCGCCCAAGTGGACAGAGTGGCCCTGCTGATTCAACAGTGCAT 2374
755 QY tAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775
2375 Db GGCCTATGAGCGGTCCAGAGGCCCTCTCTCGAGCCGCTATTCTGTGACCTCAATAGCCT 2434
775 QY uIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 795
2435 Db CATCTCTTCAGACTATGAGTCTCTCTCAGACCCACACCTGGTGGCCCTGGACCTCTGTGA 2494
795 QY pGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGl 815
2495 Db TGGGTGTGG---AATGGTGGCCAGCTCTATGCTGCTGCAAGACCCACAGTCTTCGAGGA 2551
815 QY uArgHisLeuLysTrpIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCy 835
2552 Db GAGACCTCTAAGTACATCTCACAGCTGGGAAGGGCAACTTTGGCAGCGTGGAGCTGTG 2611
835 QY sArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHi 855
2612 Db CCGCTATGACCCGCTAGCGGACATAACAGGTGCCCTGGTGGCCGCTGAAACAGCTGCAGCA 2671
855 QY sSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHi 875
2672 Db CAGCGGGCCAGACCAGCAGAGGAGCTTTTCAGCGGAGATTTCAGATCTCTCAAGCACTGCA 2731
875 QY sSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuAr 895
2732 Db CAGTGAATTCATTGTCAAGTATCGTGGTGTTCAGCTATGCGCCCGCGCCGAGAGCTGCG 2791
895 QY gLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGl 915
2792 Db GCTGTCATGAGTACCTGCGCCAGCGGCTGTCTGGCGACTCTTCGACGCGGACCCGCGC 2851
915 QY y---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTy 934
2852 Db CGCGCTCGATGCCAGCGCCCTCTCTCTATTCTTCGAGATTCTGCAAGGCGCATGGAGTA 2911
934 QY rLeuGlyAlaArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSe 954
2912 Db CCTGGGCTCCCGCGCTGCGTGCACCGGACCTGGCCGCGCCGAAACATCTCTCGTGGAGAG 2971
954 QY rGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAs 974
2972 Db CGAGGCACAGTCAAGATCGCTGACTTCGGCTTAGCTGCTGCTGCTTGAACAAGA 3031

QY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluar 1037
 Db 3150 ATACATTGAGAGAGTAAAGTCCACGCGGAATTTATCGGTATGTTGGCAATGACAA 3209
 QY 1037 gGluGlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuPr 1056
 Db 3210 ACAAGGACAGATGCTGTTCCATTTGATAGAACCTTTGAAGAAATGGAAGATTACC 3269
 QY 1056 oProProProCysProThrGluValGlnGluLeuMetGlnLeuLeuCysTrpAlaProGl 1076
 Db 3270 AGACACAGATGGATGCCAGATGATCTATATGATCATGACAGAAATGCTGGAACAATAA 3329
 QY 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093
 Db 3330 TGTAAATCAACGCCCTCTTTAGGATCTAGCTTTTCGAGTGGATCAATA 3381
 RESULT 5
 ACC46252
 ID ACC46252 standard; cDNA; 6037 BP.
 AC
 CC
 AT
 DT 02-JUN-2003 (first entry)
 XX
 DE Human dithp intracellular signalling protein-encoding cDNA.
 KW Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
 KW cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
 KW neurological disorder; gastrointestinal disorder; transport disorder;
 KW connective tissue disorder; drug screening; proteome analysis;
 KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
 KW disease model; toxicological testing; transcript imaging;
 KW intracellular signalling; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200297031-A2.
 XX
 XX 05-DEC-2002.
 XX
 XX 27-MAR-2002; 2002WO-US10056.
 XX
 XX 28-MAR-2001; 2001US-279619P.
 XX 29-MAR-2001; 2001US-280067P.
 XX 29-MAR-2001; 2001US-280068P.
 XX 16-MAY-2001; 2001US-291280P.
 XX 17-MAY-2001; 2001US-291829P.
 XX 17-MAY-2001; 2001US-291849P.
 XX 19-JUN-2001; 2001US-299428P.
 XX 20-JUN-2001; 2001US-299776P.
 XX 20-JUN-2001; 2001US-300001P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
 PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin BH;
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
 XX
 DR WPI; 2003-129518/12.
 DR P-FSDB; ABR41310.
 XX
 XX Novel human diagnostic and therapeutic polypeptide useful for
 PT identifying test compound which specifically binds to a polypeptide
 PT encoded by human diagnostic and therapeutic polynucleotide, and to
 PT induce antibodies
 XX
 PS Claim 2; SEQ ID No 173; 591pp; English.
 XX
 XX The invention relates to novel human diagnostic and therapeutic

CC polynucleotides designated dithp (ACC46080-ACC46749) and to their
 CC encoded proteins (DITHP; ABR41136-ABR41812). The invention also relates
 CC to polynucleotide sequences at least 90% identical to the dithp cDNA
 CC sequences of the invention; recombinant vectors, host cells and
 CC transgenic organisms comprising a dithp nucleic acid sequence; the
 CC recombinant production of DITHP proteins; antibodies specific for DITHP
 CC proteins; microarrays comprising dithp nucleic acid sequences; methods
 CC of detecting dithp nucleotide and protein sequences; methods of screening
 CC for compounds which specifically bind a DITHP protein; and methods of
 CC assessing the toxicity of test compounds using a dithp hybridisation
 CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
 CC diagnosis of a wide variety of conditions including cancer and other cell
 CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
 CC viral, fungal or parasitic infections; gastrointestinal disorders; metabolic
 CC disorders; neurological disorders; gastrointestinal disorders; transport
 CC disorders; and connective tissue disorders. They may also be used to
 CC screen for modulators of protein activity or gene expression. DITHP
 CC proteins can additionally be used in analysis of the proteome of a tissue
 CC or cell type and to induce antibodies. The dithp nucleic acids are
 CC additionally useful in somatic or germline gene therapy of the disorders
 CC mentioned above, as a source of antisense sequences, as a source of
 CC probes and primers, in genotyping and identification of individuals, in
 CC the generation of transgenic animal models of human disease or knock in
 CC humanised animals, in toxicological testing, and in transcript imaging.
 CC The present sequence represents a dithp cDNA encoding a DITHP protein
 CC which has intracellular signalling activity.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 6037 BP; 1959 A; 1011 C; 1206 G; 1861 T; 0 other;
 Alignment Scores:
 Pred. No.: 9 33e-191 Length: 6037
 Score: 2627.00 Matches: 534
 Percent Similarity: 66.12% Conservatives: 192
 Best Local Similarity: 48.63% Mismatches: 341
 Query Match: 44.83% Indels: 31
 DB: 25 Gaps: 14
 US-09-397-967A-16 (1-1099) x ACC46252 (1-6037)
 QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProGlu 39
 Db 512 TCCAGTTCTTCAGGT-----GTATCTTTACCATTCCTCGGAAATCTGAGGCAGA 562
 QY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLy 59
 Db 563 TTATCTGACCTTTCCATCTGGGGAGTATGTTGCAGAAAGAAATCTGTATTGCTGCTTC 622
 QY 59 sAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSe 79
 Db 623 AGCTTGTGGTATCACCTGCTGTATCATATATGCTTTTAAATGAGTGAACAGAGAAG 682
 QY 79 rCysTrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa 99
 Db 683 GATCTGGTATCCACCCACCATGCTCTCCATATAGATGATCAACAGGCATATGTACT 742
 QY 99 lTyArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr 117
 Db 743 CTACAGATAAGATTTTACTTTCTCTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 802
 QY 117 gPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137
 Db 803 GCATGGAATATCTCGAGGTGCTGAGCTCTCTCTTCTTGATGACTTTGTCTGCTTACT 862
 QY 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157
 Db 863 CTTTGCTCAGTGGCGGCATGATTTTGTGTCACGGATGGTAAAGTACCTGTGACTCATGA 922
 QY 157 sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl 177
 Db 923 AACACAGGAAGAAATGCTTTGGGATGGCAGTGTATGATGAGAAATAGCAAGAAAGA 982

3131 CATTTGTAAGTCAAGGAGTGTGCTACAGTGTCTCGCGTAATCTAAATTAATAT 3190
 898 tGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuHi 917
 3191 GGAATATTTACCATATGGAAGTTTACGAGACTATCTTCAAAAACATAAAGAACGGATAGA 3250
 917 sThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAl 937
 3251 TCACATAAAACTTCTCGAGTACACATCTCAGATATGCAAGGGATGGAGTATCTTGGTAC 3310
 937 aArgArgCysValHisArgAspLeuAlaIaArgAsnIleLeuValGluSerGluAlaHi 957
 3311 AAAAAGGTATATCCACAGGATCTGGCAACGAGAAATATATTTGGTGAGAACGAGAACAG 3370
 957 sValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTrpVa 977
 3371 AGTGAATAATGGAGATTTTGGGTTAACCAAAAGTCTTGCCACAAAGAACAAATACTATAA 3430
 977 lValArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAs 997
 3431 AGTAAAGAACCTTGGTGAAGTCCCATTTCTGGTATGCTCCAGATCATCTGACAGAGAG 3490
 997 nIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheTh 1017
 3491 CAAGTTTCTGCGCTCAGATGTTTGGAGCTTGGAGTGTCTGATGAACCTTTTCAC 3550
 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluAr 1037
 3551 ATACATTTGAGAAGAGTAAAGTTTCCACCACCGGGAATTTATGCTATGATGGCAATGACAA 3610
 1037 gGluGlyProProLeuCys---ArgLeuLeuGluLeuAlaGluGlyArgArgLeuPr 1056
 3611 ACAAGGACAGATGATGTTTCCATTTGATAGAACTTTTGAAGATATATGGAAGATTTACC 3670
 1056 oProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGl 1076
 3671 AAGACACAGATGATGCCAGATGAGATCTATATGATCATGACAGAAATGCTGGAACAATAA 3730
 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093
 3731 TGTAATCAACGCCCTCTTTTAGGGATCTAGCTCTTTCGAGTGGATCAAAATA 3782
 RESULT 6
 AAX80971
 ID AAX80971 standard; cDNA; 4482 BP.
 XX
 AC AAX80971;
 XX
 DT 03-SEP-1999 (first entry)
 XX
 DE Human JAK2 kinase encoding cDNA.
 XX
 KW JAK2 kinase; arteriosclerosis; asthma; bronchitis; emphysema; psoriasis;
 KW inflammatory bowel disease; inflammation; osteoarthritis; oncogenesis;
 KW rheumatoid arthritis; septic shock; systemic lupus erythematosus;
 KW leukaemia; human; ss.
 XX
 OS Homo sapiens.
 XX
 XX US5914393-A.
 XX
 XX 22-JUN-1999.
 XX
 XX 05-DEC-1995; 95US-0567508.
 XX
 XX 05-DEC-1995; 95US-0567508.
 XX
 XX (INCY-) INCYTE PHARM INC.
 XX
 XX Coleman R, Stuart SG;
 XX
 XX WPI; 1999-384188/32.
 DR

DR P-PSDB; AAY21698.
 XX Polypeptides and amino acids useful for modulating human jak2 kinase
 PT activity
 XX
 XX Claim 2; Columns 23-28; 37pp; English.
 CC This cDNA encodes a human JAK2 kinase polypeptide. Host cells transformed
 CC with recombinant jak2 kinase nucleic acid are used for the recombinant
 CC production of the protein. Purified JAK2 may be used to produce
 CC antibodies or identify antagonists or inhibitors of JAK2. JAK2, anti-JAK2
 CC antibodies and JAK2 antagonists or inhibitors may be used to treat,
 CC prevent or diagnose conditions associated with altered or uncontrolled
 CC jak2 expression, e.g. arteriosclerosis, asthma, bronchitis, emphysema,
 CC inflammatory bowel disease, inflammation, leukemia, oncogenesis,
 CC osteoarthritis, psoriasis, rheumatoid arthritis, septic shock and
 CC systemic lupus erythematosus.
 XX
 XX Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other;
 SQ
 Alignment Scores:
 Pred. No.: 1.8e-190 Length: 4482
 Score: 2621.00 Matches: 532
 Percent Similarity: 66.03% Conservative: 193
 Best Local Similarity: 48.45% Mismatches: 342
 Query Match: 44.73% Indels: 31
 DB: 20 Gaps: 14
 US-09-397-967a-16 (1-1099) x AAX80971 (1-4482)
 QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProGl 39
 Db 447 TCCAGTTCTTCCAGT-----GTATCTTTTACCATTCCTTGGGAAATCTTGAGGCGAGA 497
 QY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLy 59
 Db 498 TTAATCTACCTTCCATCTCGGGAGATGTTGGGAAAGAAATCTGTAATGCTGCTCTTCAA 557
 QY 59 sAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSe 79
 Db 558 AGCTTGTGTATCACACTGTGATCATATAATATGCTTTTAAATGAGTGAACAGAGAAG 617
 QY 79 rCysTrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa 99
 Db 618 GATCTGCTATCCACCAACCATGCTTCCATATAGATGAGTCAACACAGGCATATGTACT 677
 QY 99 lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr 117
 Db 678 CTACAGAAATAGATTTTACTTTCTCGTGTGATGTCAGTGGCAGCAACAGCCTATCG 737
 QY 117 gpheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137
 Db 738 GCATGGAATATCTCGAGGTGCTGAAGCTCCTCTTCTTGATGACTTTTGTGCTGTCTACCT 797
 QY 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157
 Db 798 CTTTGTCTCAGTGGCGCATGATTTTGTGTCATGGATGGATAAAAGTACCTGTGATCATGA 857
 QY 157 sGluGlyGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl 177
 Db 858 AACACAGGAAGAATGCTTGGGATGACAGTGTGTATGATGATGAGATAGCAAGAGAAA 917
 QY 177 nAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSe 197
 Db 918 CGATCAAAACCCACTGGCCATCTATACTCTATCAGCTACAAGACATTTCTTACCACAATG 977
 QY 197 rLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrVa 217
 Db 978 TATTGCGAAAGATCCAAAGACTATCATATTTTCAAGGAGCGGAAATAGGTACAGATT 1037
 QY 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLy 237
 Db 1038 TCGCAGATTTTATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAAT 1097

237 sYrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValG1 257
1098 GTATCTTATAAACTCGAAACTCTGCAGTCTGCTCTACACAGAGAAATTTGAAGTAA 1157
257 yLeuProGlyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAs 273
1158 AGAACCTGGAAGTGGCTCTTCAGCTGAGGAGATTTTTCACCAACCATATATAAATCTGAAA 1217
273 pAsnGlyIleProTyrSerSer-----AsnAspGluLe 284
1218 CGGTGAATTCAGTGTGAAGAGGGAACATAAAGAAAGTGAACACTGACAGAACAGGA 1277
284 uPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProAr 304
1278 TTTPACAGTTATATTCGATTTTCTTAATATTTATGATGTCAGTATTAAGCAACCAACCA 1337
304 gValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLe 324
1338 AGAGGGT---TCAATGAAAGCCGAGTTGTAACTATCCATAAGCAAGATCGTAAATAATCT 1394
324 uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTy 344
1395 GGAATGGAATTCAGTCTAATAGGGAAGCTTTGTCTTCGTGTCATTAATTCATGGATA 1454
344 rPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLe 364
1455 TTATAGATTAACTGCAGATGCACATCATTTACCTCTGTAAAGAGTAGCACCTCCAGCCGT 1514
364 uLeuGluGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisTy 384
1515 GCTTGAATAATATACAAAGCAACTGTCTATGCCCAATTTTCATGATGATTTTGGCATTAGTAA 1574
384 sLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTy 404
1575 ACTGAAGAAAGCAGGTAATCAGACTGGCTGATGATGATGATGATGATGATGATGATGATGAT 1634
404 rAspSerPheLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCy 424
1635 TAATAAATATTTTGGACTTTTCTGTCGCGAGGAAGATGTCATTAATATAACACGTG 1694
424 sLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAr 444
1695 TTTGATTACAAAATAATGCAATGCAAGAGTACAACTCAGTGGGACAAAGAACTTCAG 1754
444 gSerLeuArgGluLeuAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAl 464
1755 CAGTCTTAAAGATCTTTTGAATTTTACCGAGATGGAAACTGTTCTGCTCAGACAAATATAAT 1814
464 aLeuTyrLeuThrSerCysAlaProArgProLysGluLysSerAsnLeuIleValVa 484
1815 TTTCCAGTTTACTAAATGCTGTCCTCCCAAGCCAAAGATAATCAAACTTCTAGTCTT 1874
484 lArg---ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysCysAlaLe 503
1875 CAGAACGAATGGTGTCTGTGATGATACCAACCTCACCACATTTACAGAGCCCTACTCATAT 1934
503 uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGl 523
1935 GAACCAATGGTGTTCACAAAATCAGAAATCAAGATTTGATATTAATGAAGCCCTTGG 1994
523 yHisGlySerPheThrLysIlePheArgGlySerArgGluValValAlaAsp---GlyGl 542
1995 CCAAGGCATTTTACAAAGATTTTAAAGCGCTGCAAGAGAGATGAGGACTACCGTCA 2054
542 uThrHisAspSerGluValLeuLysValMetAspSerArgHisArgAsnCysMetGl 562
2055 ACTGCATGAAACAGAGATCTTTTAAAGTCTTGGATAAAGCACACAGCAACTATTCAGA 2114
562 uSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLe 582
2115 GTCTTTCTTGAAGCAGCAAGTATGATGAGCAAGCTTTCTCACAAGCATTTTGGTTTAA 2174

582 uHisGlyValCysMetAlaGlyAsp---SerIleMetValGlnGluPheValTyrLeuGl 601
2175 TTATGAGATATGTCGTGGAGACGAGAAATATCTGGITTCAGAGTITGTAAATTTGG 2234
601 yAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLysLeuGl 621
2235 ATCACTAGATACATATCTGAAGAAGAAATAAAATTTGTATAAATATATATTATGGAACCTGA 2294
621 nValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGl 641
2295 AGTTGTAAACAGTTGGCATGGCCATGCTATTTCTAGAGAAAAACACCTTTATTCATGG 2354
641 yAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGly---AspGlyAsnPr 659
2355 GAATGATGTGCCCAAAATATTTCTGCTTATCAGAGAAGACAGGAAAGACAGAAATCC 2414
659 oProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuLeuMetLe 679
2415 TCCTTTTCATCAAACTTAGTGATCTCGCATTAGTATTACAGTTTTCGCAAGACATTTCT 2474
679 uThrAspArgIleProTyrValAlaProGluCysLeuGlnGluAlaGlnThrLeuCyste 699
2475 TCAGGAGATATACCATGGGTACCACTGAATGATTTGAAAAATCTCTAAAAATTTAAATTT 2534
699 uGluAlaAspLysTyrGlyPheGlyAlaThrThrTyrGluValPheGlnArgGlyProAl 719
2535 GGCAACAGACAATGGAGTTTGGTACCACCTTTGTGGAAATCTGCAGTGGAGGAGATAA 2594
719 aHisIleThrSerLeuGluProAlaLysLysLysLysPheTyrGluAspGlnGlyGlnLe 739
2595 ACCTCPAAGTCTCTGGAATCTCAAGAAGACTACAATTTATGAAGATAGGCATCAGCT 2654
739 uProAlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPr 759
2655 TCTGCACCAAGTGGGCAGAAATTAGCAAACTTATAAATAATTTGATGGATTTGAACC 2714
759 oGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs 779
2715 AGATTTTCAGGCTCTCTTTTCAGAGCCATCATACAGATCTTAAACAGTTTGTACTCCAGA 2774
779 pTyrGluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCy 798
2775 TTATGACTATTAAACAGAA---AATGACATGTTTCCAAATATAGAGATAGTGCCTTGG 2831
798 sValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLe 818
2832 GTTTCTGTGTGCC-----TTTGAAGACCGGATCTCTACACAGTTTGAAGAGAGACATTT 2885
818 uLysTyrIleSerLeuLeuGlyLysGlyAsnPhedGlySerValGluLeuCysArgTyrAs 838
2886 GAAATTTTTCAGCAACTTTGGCAAGGGTAATTTTGGAGTGTGGAGATGTCCGGTATGA 2945
838 pProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPr 858
2946 CCTCTCAGGACACACTGGGAGGTGGTCTGCTGTAATAAAGCTTTCAGATAGTACTGA 3005
858 oAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPh 878
3006 AGAGCACCTTAAGACACTTTGAAAGGAAATTTGAAATCTGAAATCCCTACACATGACAA 3065
878 eIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMe 898
3066 CATTGTAAGTACAAAGGAGTGTGCTACAGTGTCTGCTCGCGTAAATCTTAAATAATTAAT 3125
898 tGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuHi 917
3126 GGAATATTTTCCATATGAAAGTTTACGACATCTCTTCAAAAACATAAAGACGACATAGA 3185
917 sThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAl 937
3186 TCACATAAAACTTCTGCAGTACACATCTCAGATATGCAAGGGTATGGAGTATCTTGGTAC 3245
937 aArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHi 957

```

DB 3246 AAAAGGTATATCCAGGGATCTGGACGAGAAATATATTGTTGGTGGAGAACGAGACAG 3305
QY 957 sValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTrpVa 977
DB 3306 AGTTAAATTTGGTGATTTTGGTTTAAACCAAGTCTTGGCCACAGACAAAGAAATCTATATA 3365
QY 977 lValArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAs 997
DB 3366 AGTAAAGAACCTGGTGAAGTCCCATATTTCTGGTATGCTCCAGATCACTGACAGAGAG 3425
QY 997 nIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheTh 1017
DB 3426 CAAGTTTCTGTCGCTTCAGATGTTTGGAGCTTTGGAGTGGTCTGTATGAACCTTTTCAC 3485
QY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluAr 1037
DB 3486 ATACATTGAGAAGAGTAAAGTCCACGCGGAATTTATCGTATGATTGGCAATGACAA 3545
QY 1037 gGluGlyProProLeuCys---ArgLeuLeuGluLeuAlaGluGlyArgArgLeuPr 1056
DB 3546 ACAGGACAGATGATCGTGTTCATTTGATAGAACTTTTGAGAAATATGGAAGATTACC 3605
QY 1056 oProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGl 1076
DB 3606 AAGACACAGATGGATGCCAGATGAGATCTATATGATCATGACAGAAATGCTGGMAACAATA 3665
QY 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093
DB 3666 TGTAATCAAGCCCTCTTTAGGGATCTAGCTCTTGGAGTGGATCAATA 3717

```

```

RESULT 7
AAZ58947
ID AAZ58947 standard; cDNA; 4482 BP.
XX AC AAZ58947;
XX DT 03-MAY-2000 (first entry)
XX DE Human JAK2 kinase (HJAK2) encoding cDNA.
XX KW Janus family nonreceptor protein-tyrosine kinase-2; JAK2; HJAK2; human;
XX KW signal transduction; arteriosclerosis; asthma; bronchitis; emphysema;
XX KW inflammatory bowel disease; leukemia; oncogenesis; osteoarthritis;
XX KW psoriasis; rheumatoid arthritis; systemic lupus erythematosus;
XX KW cytosolic; osteopathic; dermatological; antibacterial; septic shock;
XX KW immunosuppressive; ss.
XX OS Homo sapiens.
XX PN US6019966-A.
XX PD 01-FEB-2000.
XX PF 19-NOV-1998; 98US-0196480.
XX PR 05-DEC-1995; 95US-0567508.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Stuart SG, Coleman R;
XX DR WPI, 2000-146859/13.
XX DR P-PSDB; AAY77552.
XX PT Human Janus family nonreceptor protein-tyrosine kinases useful as
XX PT diagnostic reagents and for preventing, diagnosing and treating
XX PT diseases such as arteriosclerosis, asthma and leukemia -
XX FS Example 5; Fig 1A-F; 33pp; English.
XX CC This cDNA encodes a human Janus family nonreceptor protein-tyrosine
XX CC kinase-2 (JAK2) polypeptide (HJAK2). The JAK2 polypeptides may be used

```

as diagnostic reagents as they react with a range of target proteins including growth hormone, prolactin, erythropoietin and cytokine receptors. They may also be used for the production of antibodies specific for JAK2, which may be used to inhibit its activity and prevent or treat disorders associated with over expression of JAK2. Conversely, the JAK2 polypeptide may be administered to supplement the patients own production and counter mutations that may lead to the expression of an inactive enzyme. The protein may also be used to screen candidate reagents for modulators of JAK2 function. The antagonists and antibodies bind to the JAK2 protein and prevent the transfer of high energy phosphate molecules, therefore blocking signal transduction. Disorders that may be treated by administration of JAK2 polypeptides, anti-JAK2 antibodies and the agonists and/or antagonists, include arteriosclerosis, asthma, bronchitis, emphysema, inflammatory bowel disease, leukemia, oncogenesis, osteoarthritis, psoriasis, rheumatoid arthritis, septic shock and systemic lupus erythematosus.

SQ Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other;

Alignment Scores:
 Pred. No.: 1.8e-190 Length: 4482
 Score: 2621.00 Matches: 532
 Percent Similarity: 66.03% Conservative: 193
 Best Local Similarity: 48.45% Mismatches: 342
 Query Match: 44.73% Indels: 31
 DB: 21 Gaps: 14

US-09-397-967A-16 (1-1099) x AAZ58947 (1-4482)

```

QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProGl 39
DB 447 TCCAGTTCTTCAGGT-----GTATCTTTACCATTCCTCCCTGGGAAATCTGGAGGAGA 497
QY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLy 59
DB 498 TTATCTGACCTTTCCATCTGGGAGTATGTTGGAGAGAAATCTGTATGCTGCTCTTAA 557
QY 59 sAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSe 79
DB 558 AGCTTGTGGTATCACACCTGTATCATATAATATCTTTGCTTTAATGATGAGTGAACAGAAAG 617
QY 79 rCysTrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa 99
DB 618 GATCTGTATTCACCCCAACCATGCTTCCATATAGATGAGTCAACAGGCAATATGTAAT 677
QY 99 lTyrArgLeuArgPheTyrPheProAspTyrPhe-----GlyLeuGluThrCysHisAr 117
DB 678 CTACAGATATAGATTTTACTTTCTCTGTTGATGTCAGTGGCAGCAACAGAGCTATCG 737
QY 117 gPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137
DB 738 GCATGGAATATCTCGAGGTGCTGAAGCTCTCTTCTTGTATGATGACATTTGTCTATCTTACT 797
QY 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157
DB 798 CTTTGTCTCAGTGGCGGCATGATTTTGTGCATGGATGGAATAAAGTACTGTGACATCATGA 857
QY 157 sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl 177
DB 858 AACACAGGAAGATGCTTGGATGACAGATCTTAGATATGATGAGAAATAGCAAGAAAA 917
QY 177 nAlaGlnArgProGlyGluLeuLysThrValSerTyrLysAlaCysLeuProProSe 197
DB 918 CGATCAAAACCCACTGGCCATCTATACTCTATCAGTCTACAGACATTTCTTACCACAATG 977
QY 197 rLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgGlyLeuArgThrVa 217
DB 978 TATTGACCAAGATCCCAAGTATCATATTTTGAACAAGAGCGCAATAGTACAGATT 1037
QY 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLy 237
DB 1038 TCGCAGATTATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAG 1097

```

QY 237 sTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValG1 257
DB 1098 GTATCTTATAAATCTGAAACTCTGAGTCTGCTTCTACACAGAGAAATTTGAAGTAA 1157
QY 257 yLeuProGlyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAs 273
DB 1158 AGAACCTGGAAGTGGTCTCTCAGGTGAGGAGATTTTTCACCATTAATAAATCTGAA 1217
QY 273 pAsnGlyIleProTTPSerSer-----AsnAspGluLe 284
DB 1218 CGTGGAAATTCAGTGTCAAGGGAACATAAAGAGTGAGACTGACAGAGCAGA 1277
QY 284 uPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProAr 304
DB 1278 TTACAGTTATATTCGATTTTCTTAATATTATTGATGTCAGTATTAAAGCAAGCAACCA 1337
QY 304 gValGlyProAlaGlyGluHisArgLeuValThrValArgMetAspGlyHisIleLe 324
DB 1338 AGAGGGT---TCAATGAAAGCCGAGTTGTACTATCCATAGCAAGATGGTAAATAATCT 1394
QY 324 uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTy 344
DB 1395 GGAATTCGACTGATCATTAAGGAGACCTTTGCTTCGTGTCATTAATTTGATGATA 1454
QY 344 rPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLe 364
DB 1455 TTATAGATTAACTGCAGATGCATCATTTACCTCTGTAAAGAGTAGCACCTCCAGCCGT 1514
QY 364 uLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisIy 384
DB 1515 GCTTGAATAATATACAAAGCAACTGTCTGCGCCCAATTTTCGATGGATTTTGCATTTAGTAA 1574
QY 384 sLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTy 404
DB 1575 ACTGAGAAAGCAGAGTAACTGAGTGGAGCTGTATGATCTTCGATCGCTCTTAAGGACTT 1634
QY 404 rAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCy 424
DB 1635 TAATAAATATTTTGGCTTTGCTGTCGCGAGAGAAATGTCATTGAATATAACACTG 1694
QY 424 sLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAr 444
DB 1695 TTTGATTACAAAATAATAGAAATGAAGAGTACAACTCAGTGGGCAAAAGAACTTCAG 1754
QY 444 sSerLeuArgGluLeuLeuAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAl 464
DB 1755 CAGCTTAAAGATCTTTGAATTTTACAGATGAAACTGTTCTGCTCAGACATATTAAT 1814
QY 464 aLeuTyrLeuThrSerCysAlaProArgProLysGluLysSerAsnLeuIleValVa 484
DB 1815 TTTCAGTTTACTAAATGCTGTCCCAAGCCAAAGATATAATCAAACTTCTAGTCTT 1874
QY 484 lArg---ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAlaLe 503
DB 1875 CAGAACGAATCGTGTCTGATGTACCAACCTCCACCAATTCAGAGGCTCTCATAT 1934
QY 503 uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTyrHisGluAsnLeuG1 523
DB 1935 GAACCAATGTGTTCACAAATTCAGAAATGAGATTTGATATTATTAATGAAGCCITGG 1994
QY 523 yHisGlySerPheThrIlePheArgGlySerArgArgGluValValAsp---GlyG1 542
DB 1995 CCAAGGCATTTTACAAAGATTTTAAAGCGGTACGAAGAGAGTAGGAGACTACGGTCA 2054
QY 542 uThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetG1 562
DB 2055 ACTGCATGAAACAGAAAGTCTTTTAAAGATTTCTGGATAAAGCACACAGCAACTATTCA 2114
QY 562 uSerPheLeuGluAlaAlaSerLeuMetSerClnValSerTyrProHisLeuValLeuLe 582
DB 2115 GTCCTTCTTTGAGCAGCAAGTATGAGCAAGCTTTCTCACAGCAATTTGGTTTTAA 2174
QY 582 uHisGlyValCysMetAlaGlyAsp---SerIleMetValGlnGluPheValTyrLeuG1 601

DB 2175 TTATGGAGTATGCTGCTGGAGACAGAAATATCTGGTTCAGGAGTTTGTAAATTTGG 2234
QY 601 yAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrLysLeuG1 621
DB 2235 ATCCTAGATACATATCTCAAAAAGAAATAAAATTTGTATAATATATTATGGAACCTTGA 2294
QY 621 nValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHisG1 641
DB 2295 AGTTCTAAACAGTTGGCATGGCCATGTCATTTCTAGAAGAAACACCCCTTATTCTATGG 2354
QY 641 yAsnValSerAlaAlaGlyLysValLeuLeuAlaArgGluGlyGly-----AspGlyAsnPr 659
DB 2355 GAATGATGTGCCAAAATAATTCCTTATCAGAGAAGACAGACAGAAACAGGAATCC 2414
QY 659 oProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLe 679
DB 2415 TCCTTTTCATCAAACTTAGTATCTCTGGCATTAGTATTACAGTTTTTGCCTGCAAGACATTTCT 2474
QY 679 uThrAspArgIleProTTPValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLe 699
DB 2475 TCAGGAGAGATACCATGGGTACCCCTGAATGCATTGAAATTCCTAAATAATTTAAATTT 2534
QY 699 uGluAlaAspLysTyrGlyPheGlyValAlaThrThrTyrGluValPheGlnArgGlyProAl 719
DB 2535 GGCACACAGCAAAATGGAGTTTGGTACCACCTTGTGGGAATCTCAGTGGAGGAGATAA 2594
QY 719 aHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLe 739
DB 2595 ACCTCTAAGTCTCTGGATTTCTCAAAAGAAAGTCAATTTTATGAAGATAGCATCAGCT 2654
QY 739 uProAlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPr 759
DB 2655 TCCTGCACCAAGTGGGAGAAATTAGCAAACTTATAAATAATTTGATGGATTATGAACC 2714
QY 759 oGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs 779
DB 2715 AGATTCAGGCTCTCTTTCAGAGCCATCATAGAGATCTTAAACAGTTTGTCTACTCCAGA 2774
QY 779 pTyrGluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCy 798
DB 2775 TTATCAACTATTAAACAGAA---AATGACATGTTACCAATATGAGATAGGTGCTTGGG 2831
QY 798 sValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLe 818
DB 2832 GTTTCTGCTGCC-----TTTGAAGACCGGATCTCTACACAGTTTGAAGAGAGACATTT 2885
QY 818 uLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAs 838
DB 2886 GAAATTTCTACAGCAACTTGGCAAGGTAAATTTGGGAGTGTGGAGATGTCGCGGTATGA 2945
QY 838 pProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPr 858
DB 2946 CCCTCTACAGGACAACTGGGGAGGTGCTCGCTGTAATAAAGCTTCAGCATAGTACTGA 3005
QY 858 oAspGlnClnArgAspPheGlnArgGluIleClnIleLeuLysAlaLeuHisSerAspPh 878
DB 3006 AGAGCACCTTAAGACACTTTGAAGGGAAATTTGAATCTGAAATCCCTACAGCATGACAA 3065
QY 878 eIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMe 898
DB 3066 CATTTGTAAGTACAAAGGAGTGTCTACAGTGTGGTGGCGGTAAATCTAAATAATTAATAT 3125
QY 898 tGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHi 917
DB 3126 GGAATATTTACCATATGGAAGTTTACGAGACTATCTTCAAAAAACATAAAGAACCGATAGA 3185
QY 917 sThrAspArgLeuLeuPheAlaTyrPtnIleCysLysGlyMetGluTyrLeuGlyAl 937
DB 3186 TCACATAAACTCTCGAGTACACATCTCAGATATGCAAGGGTATGAGATATCTTGGTAC 3245
QY 937 aArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHi 957

3246 AAAAGGTTATATCCACAGGATCTGGCAACGAGAAATATATTTGGTGGAGAACGAGAACAG 3305
 QY 957 sVallysHleAlaAspPheGlyLeuAlaLysLeuProLeuGlyLysAspTyrVa 977
 Db 3306 AGTTAAATTTGGGATTTGGGTTAACCAAGTCTTGGCCACAGACAAAGATCTATATA 3365
 QY 977 lValArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAs 997
 Db 3366 AGTAAAGAACCTGGTGAAGTCCCATATTTCTGGTATGCTCCAGATCACTGACAGAG 3425
 QY 997 nilePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPhe 1017
 Db 3426 CAAGTTTCTGTGGCTCAGATGTTGGAGCTTTGGAGTGTCTGTATGAATTTTAC 3485
 QY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGlu 1037
 Db 3486 ATACATTGAGAGAGTAAAGTCCACGAGCGGAATTTATGCGTATGTTGGCAATGACAA 3545
 QY 1037 sGluGlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeu 1056
 Db 3546 ACAAGACAGATGATCGTGTCTTCATTGTATGATGATTTTGAAGAATAATGAGATACC 3605
 QY 1056 oProProProThrCysProThrGluValGlnGluLeuMetGlnLeuMetGlnLeuPro 1076
 Db 3606 AAGACAGATGGATGCCAGATGAGATGATATATGATCATGACAGAAATGCTGGAACAATA 3665
 QY 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093
 Db 3666 TGTAATCAACGCCCTCTTTAGGGATCTAGCTCTTCGAGTGGATCAATA 3717

RESULT 8

AAD24311
 ID AAD24311 standard; DNA; 5117 BP.
 AC AAD24311;
 XX
 DT 07-MAR-2002 (first entry)
 DE Human Jak2 (Janus kinase) DNA.
 XX
 KW Human; SOCS; suppressor of cytokine signalling; autoimmune disorder; Jak;
 KW Janus kinase; signal transducer and activator of transcription; Stat;
 KW transplant acceptance; systemic lupus erythematosus; glomerulonephritis;
 KW rheumatoid arthritis; Wegener's granulomatosis; chronic active hepatitis;
 KW atopic dermatitis; multiple sclerosis; myasthenia gravis; vasculitis;
 KW Crohn's disease; haemolytic anaemia; nephrotic syndrome; dermatological;
 KW diabetes mellitus; thyroiditis; inflammatory bowel disease; nephrotropic;
 KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
 KW hepatotropic; ds.
 XX
 OS Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT CDS 478..3876
 FT /tag= a
 FT /product= "Human Jak2 protein"
 XX
 PN WO200179555-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 13-APR-2001; 2001WO-US12131.
 XX
 PR 14-APR-2000; 2000US-0549654.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 PI Hancock WW, Ozkaynak B;
 XX
 DR WFI; 2002-034368/04.
 DR P-PSDB; AAE15179.
 XX
 PT Monitoring transplant acceptance or autoimmune disease, useful e.g. for

PT assessing therapy, comprises measuring levels of Stat or their
 inhibitors
 XX
 PS Example; Fig 8; 218pp; English.
 XX
 CC The present invention relates to a method for monitoring acceptance of a
 CC transplant or an autoimmune disease in a mammal. The method comprising
 CC determining the amount of at least one of Stat4 (signal transducer and
 CC activator of transcription), Stat6, SOCS1 (suppressor of cytokine
 CC signalling) or SOCS3 mRNA or protein in a sample of the transplant, taken
 CC from the host or an affected tissue sample. Stats are activated by
 CC receptor-associated Janus kinases (Jaks) which include Jak1, Jak2, Tyk2,
 CC Jak3. The method is used to determine whether acceptance of a transplant
 CC has been induced or to determine if autoimmune disorders (systemic
 CC lupus erythematosus, glomerulonephritis, rheumatoid arthritis, Wegener's
 CC granulomatosis, chronic active hepatitis, atopic dermatitis, multiple
 CC sclerosis, myasthenia gravis, haemolytic anaemia, nephrotic syndrome,
 CC thyroiditis, diabetes mellitus, Crohn's disease, inflammatory bowel
 CC disease and vasculitis) are being treated successfully and may be used
 CC to adjust treatment regimes. The present sequence is human Jak2 DNA.
 XX
 SQ Sequence 5117 BP; 1623 A; 927 C; 1103 G; 1464 T; 0 other;

Alignment Scores:

Pred. No.: 2,15e-190 Length: 5117
 Score: 2621.00 Matches: 533
 Percent Similarity: 66.12% Conservative: 193
 Best Local Similarity: 48.54% Mismatches: 341
 Query Match: 44.73% Indels: 31
 DB: 24 Gaps: 14

US-09-397-967A-16 (1-1099) x AAD24311 (1-5117)

QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProGl 39
 Db 585 TCCAGTCTCTTCAGGT-----GTATCTTTACCATTCCTCTGGGAATCTGAGGAGA 635
 QY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaL 59
 Db 636 TTATCTGACCTTCCATCTGGGAGTATGTTGAGAGAAATCTGTATTCTGCTTCTAA 695
 QY 59 sAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPhe 79
 Db 696 AGCTTGTGTATCACACCTGTATCATATATGTTTGTCTTATGATGAGTGAACAGAAAG 755
 QY 79 rCysTrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa 99
 Db 756 GATCTGTATCCACCAACCATGCTCTTCATATAGATGATCAACAGGCATATGTACT 815
 QY 99 lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr 117
 Db 816 CTACAGATPAAGATTTTACTTTCTCTGTATTGAGTGGCAGCAACAGACCTATCG 875
 QY 117 gPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137
 Db 876 GCATGGAATATCTCGAGGTGCTGAAGCTCCTCTTCTTGTAGTACTTTTCTCATCTTACCT 935
 QY 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157
 Db 936 CTTTGTCTCAGTGGCGCATGATTTTGTGACGATGATGATGATGATGATGATGATGATG 995
 QY 157 sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl 177
 Db 996 AACACAGAGAAGATGCTTGGAGTGGCAGTGTGTAGATATGATGAGATGAGTGAAGAAA 1055
 QY 177 nAlaGlnArgProGlyGluLeuLysThrValSerTyrIlyslalCysLeuProProSe 197
 Db 1056 CGATCAAAACCCCACTGGCCATCTATAACTCTATCAGCTACAAGACATCTTACCAAAATG 1115
 QY 197 rLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgIleArgArgThrVa 217
 Db 1116 TATTTCGAGAAAGATCCAGACTATCATATTTTTCACAGGAGGAGGATAGTACAGATT 1175

QY	761	ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr	780	AC	AAC66244;
Db	2473	TTTCAGCCCTGCTTTCAGAGCTGTCCGATGCTTTAAACAGCCTGTTTACTCAGATTAT	2532	XX	19-FEB-2001 (first entry)
QY	781	GluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCysVal	799	XX	Jak2 polynucleotide sequence.
Db	2533	GAACACTACTACAGAA---AATGACATGCTACCAACATGAGATAGTAGGCTT	2589	XX	Jak3; kinase; cytokine; cellular response; inhibition; jak2; ss;
QY	800	AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLys	819	XX	cell proliferation; erythrocytosis.
Db	2590	TCTGTGCT-----TTTGAAGACAGGAGCCCTACACAGTTTGAAGAGACACTTGAAG	2643	XX	Unidentified.
QY	820	TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspPro	839	XX	US6136595-A.
Db	2644	TTTCTACAGAGCTTGGCAAGGTAATCTCGGAGTGTGGAGATGCGCCCTATGACCCG	2703	XX	24-OCT-2000.
QY	840	LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp	859	XX	18-JUN-1996; 96US-0665574.
Db	2704	CTGCAGACAACACTGGCGAGGTGGTGGCTGTGAAGAACTCCAGACACAGCACTGAAGAG	2763	XX	29-JUL-1994; 94US-0282012.
QY	860	GlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIle	879	XX	29-JUL-1993; 93US-0097997.
Db	2764	CACCTCGAGACTTTCAGAGGAGATCGAGATCTGAAATCTTGCAGCATGACAACATC	2823	XX	09-SEP-1993; 93US-0118968.
QY	880	ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu	899	XX	(SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
Db	2824	GTCAAGTACAAAGGAGGTGCTCAGTGGCGGTGCGGCAACCTTAAGATTAAATATGAA	2883	XX	Silvenmoinen O, Witthuhn BA, Ihle J;
QY	900	TyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuHisThr	918	XX	WPI; 2000-686080/67.
Db	2884	TATTTACCATATGGAAGTTTACGAGACTATCTCCAAAACATAAAGAACGATGATCAC	2943	XX	P-PSDB; AAB35719.
QY	919	AspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArg	938	XX	New DNA encoding Jak3 kinase is useful as cytokine regulator for
Db	2944	AAAAAATCTTCAATACACATCTCAGATATGCAAGGCGATGGAATATCTTGGTACAAA	3003	XX	treating cell proliferation -
QY	939	ArgCysValHisArgAspLeuAlaAArgAsnIleLeuValGluSerGluAlaHisVal	958	XX	Example 1; Fig 1; 100pp; English.
Db	3004	AGGTATATCCACAGGACCTGGCAACAGGAACATATGTGGAAAATGAGAACAGGTT	3063	XX	This invention relates to DNA encoding a murine Jak3 protein. The amino
QY	959	LysIleAlaAspPheGlyLeuAlaLysLeuProLeuGlyLysAspTyrTyrValVal	978	XX	acid sequence of the Jak3 protein is given in AAB35715. The Jak family
Db	3064	AAATAGGAGACTTCGATTAACCAAGTCTTGGCGGAGCAAGAACTACTACAAAGTA	3123	XX	of kinases are involved in the cellular response to the binding of
QY	979	ArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnIle	998	XX	cytokines to their respective receptors. Jak3 kinase mediated activation
Db	3124	AAGGAGCCAGGGAAAGCCCATATCTGGTACGACCTGAATCCTTTCGCGGAGCAAG	3183	XX	of some cytokines through their phosphorylation in response to
QY	999	PheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr	1018	XX	CC cytokine-receptor binding. Inhibiting the activity of Jak3 kinase (at the
Db	3184	TTTTCTGTGGCTCAGATGTGTGGAGCTTTGGAGTCGTTCTATACGAACTTTTCACATAC	3243	XX	CC nucleic acid level with antisense sequences or ribozymes, or at the
QY	1019	CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGlu	1038	XX	CC protein level with antibodies, kinase inhibitors etc.) is used to treat,
Db	3244	ATCGAGAGAGTAAAGTCCACCGTGGAAATTTATGCAATGATGCAATGATAAACA	3303	XX	CC e.g. excessive cell proliferation such as erythrocytosis. Alternatively,
QY	1039	GlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProPro	1057	XX	CC the Jak3 protein and polynucleotide can be used to treat conditions
Db	3304	GGGCAATGATGTGTCTTCCATTTGATAGAGTACTGAAGAGCAACGGAAGATGCCAAGG	3363	XX	CC associated with defective Jak3 activity. The DNA sequence can be used to
QY	1058	ProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGluPro	1077	XX	CC produce recombinant Jak3 and this used to raise antibodies useful as
Db	3364	CCAGAAGATCCCATGATGATTTATGTATCATGATGACAGAGTCTGGAACAACAAATGTG	3423	XX	CC specific inhibitors or to detect or isolate Jak3 without interfering
QY	1078	HisAspArgProAlaPheAlaThrIleuSer	1087	XX	CC with its enzymatic activity. The present sequence represents cDNA
Db	3424	AGCCAGCTCTCTCTCTCTCAGGACCTTTCG	3453	XX	CC encoding a Jak2 protein used in the isolation and characterisation of the
QY				XX	CC Jak3 protein of the invention.
QY				XX	SQ Sequence 3629 BP; 1142 A; 732 C; 843 G; 912 T; 0 other;
QY				XX	Alignment Scores:
QY				XX	Pred. No.: 1,76e-190 Length: 3629
QY				XX	Score: 2619.50 Matches: 537
QY				XX	Percent Similarity: 65.50% Conservative: 190
QY				XX	Best Local Similarity: 48.38% Mismatches: 350
QY				XX	Query Match: 44.70% Indels: 33
QY				XX	DB: 21 Gaps: 15
QY				XX	US-09-397-967A-16 (1-1099) x AAC66244 (1-3629)
QY				XX	Qy 4 ProSerGluThrProLeuIleProGlnArgSerCysSerLeuSerSerSerGluAla 23
QY				XX	Db 145 CCTGTACATCAAGATGTGTATTTCTTGGAGTGCTAATCTGTGACGACATAGAGCCA 204
QY				XX	Qy 24 GlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArgLeuSerPhe 43
QY				XX	Db 205 GTCCTTCAAGTGTATCTGTACCATCTCTTGGCAAGCTGAAGAGAGATATCTGAAGTTT 264
QY				XX	Qy 44 SerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAlaCysGlyTle 63

265 CCAAGTGGAGAGTATGTTGTCAGAGAAATTTGTGGCTGCTCTTAAAGCTTGTGGTATT 324
Db
64 LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPhePro 83
Qy
325 ACGCTGTGTATCAATAATATGTTGCGTTAATAGTGAACCCGAAAGGATCTGGTACCCA 384
Db
84 ProSerHisIlePheCysIleGluValAspValAspThrGlnValLeuValTyrArgLeuArg 103
Qy
395 CCCAATCATGCTTCCCATAGACGAGTCAACAGGATGACATCTCTACAGGATAGG 444
Db
104 PheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisArgPheGlyLeuArg 121
Qy
445 TTCTACTTTCCCTCATGTTGTTAGTGGCAGCAGAGAACCTACAGATACGAGGTGCC 504
Db
122 LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141
Qy
505 COTGGGGTGAAGCTCTCTGCTGATGACTTTGTCATGCTTCTTACCTTTTGTCTCAGTGG 564
Db
142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGlu 161
Qy
565 CGGCATGATTTGTTCAAGGATGATATAAAGTACCTGTGACTCATGAACCTCAGGAAG 624
Db
162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgPro 181
Qy
625 TGTCTTGGGATGGCGGTGTAGACATGATGAGATAGCTTAAGGAGAAACACCACTCCA 684
Db
182 GlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal 201
Qy
685 CTGGCTGTCTAATCTGTCTAGCTTACAGACATTTCTTACCAAGTGGCTTCGAGCGAAG 744
Db
202 IleGlnGlyGlnAsnPheValThrArgArgIleArgThrValValLeuAlaLeu 221
Qy
745 ATCCAAGACTATCACATTTTAAACCGGAAGCAATCAGTACAGATTTTCGCAGATTCATT 804
Db
222 LeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAsp 241
Qy
805 CAGCAATTCAGTCAATGTAAAGCCACTGCCAGGAACCTTAAAGTATCTTATAAAC 864
Db
242 LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal----- 256
Qy
865 CTGMAACCTTGAGCTGCTCTTACAGAACAGTGTGAAAGTAAAGAAATCTGCAAGA 924
Db
257 GlyLeuProGlyAlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle 276
Qy
925 GGT---CCTCAGGTGAGGAGATTTTGCACACATTAATACTGTAAGACCGTGGAAAT 981
Db
277 ProTrpSerSer-----AsnAspGluLeuPheGlnThr 287
Qy
982 CAGTGGTCAAGAGGGAAACATAAGGAAAGTGAAGACACTGACAGAACAGGACGTACAGTTA 1041
Db
288 PheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyPro 307
Qy
1042 TATGTGTATTTCCCTGTATTAATGATGTCAGTATTAGCAAGCAACACG---GAATGC 1098
Db
308 AlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGlu 327
Qy
1099 TCAATGAAGTAGAATGTAACTGCTCCATAACCAAGATGGTAAAGTTTGGAGATAGAA 1158
Db
328 PheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347
Qy
1159 CTTAGCTCATTAAGAAAGAGCCCTGTGTCATTCGTGTCTAATTAATGACGGGTATTACAGACTA 1218
Db
348 IleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGluGlu 367
Qy
1219 ACTGGGATGCGCACCATTAATCTGTCAAGAGGTTGGCTCCCGAGCTGTGTCAGAAC 1278
Db
368 GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAla 387
Qy
1279 ATACACAGCAACTGCCCGCCCATATCAATGGATTTTGCAATTAGCAAACTAAGGAAG 1338
Db
388 AlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspPhe 407
Qy
1339 GCGGGTAAACAGACTGAGCTATATGTCTACGATGACGCCCTTAAGGACTTCAACAAATAC 1398
Db
408 LeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArg 427
Qy
1399 TTTCTGACCTTTGCTGTGTGAGCGAGAAAATGTCATTGAATATAAACACTGTTTGTATTACG 1458
Db
428 GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg 447
Qy
1459 AAGATGAGATGGAGAAATACAACTCAGCGGAGCTAAGAGGAACCTTCAGTAACCTTAAG 1518
Db
448 GluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeu 467
Qy
1519 GACCTTTTGAATGTGTACAGATGAAAATGTGCGCTCAGACAGTATCATCTTCAGATT 1578
Db
468 ThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIleValValArgGly 487
Qy
1579 ACCAAATGCTGCCCCCAAGCAAAAGATAAATCAAACTTCTCGTCTTCAGA----- 1632
Db
488 CysAsnProAlaProAlaProGlyCysSerProSerCysCysAla-----LeuThr 504
Qy
1633 ACAATGCTATTTCTGTATGTTTCAGATCTCACCACATACAGAGGCATAAATATGTGAAT 1692
Db
505 GlnLeuSerPheHisThrIleProThrAspSerLeuGluTyrHisGluAsnLeuGlyHis 524
Qy
1693 CAAATGCTGTTTCACAAAATCAGGAATGAAGATTTAATATTAAATCAAAAGTCTTGCCCAA 1752
Db
525 GlySerPheThrLysIlePheArgGlySerArgArgGluValValValAsp---GlyGluThr 543
Qy
1753 GGTACTTTTACAAAATTTTAAAGGTGAAGAGAGATTTGGAGATTTATGGTCAACTG 1812
Db
544 HisAspSerGluValLeuLysValMetAspSerArgHisArgAsnCysMetGluSer 563
Qy
1813 CACAAACGGAAGTTCTTTGAAAGTCTAGATAAGCACATAGGAATTAATTCAGAGTCT 1872
Db
564 PheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuHis 583
Qy
1873 TTTCTCGAAGCAGCAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1932
Db
584 GlyValCysMetAlaGly---AspSerIleMetValGlnGluPheValTyrLeuGlyAla 602
Qy
1933 GGTGTCTGTCTGTGGAGAGGAGAACATTTCTGGTTCAAGAAATTTGTAATAATTTGGATCA 1992
Db
603 IleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrLysLeuGlnVal 622
Qy
1993 CTGGATACATACCTCAGAGAGAAACAAAATTTCCATAATATATATATGAAACTTGGAGTG 2052
Db
623 ThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsn 642
Qy
2053 GCTAAGCAGTTCGGATGGCCATGATTTCTAGAGAGAAATCCCTTATTTCATGGGAAT 2112
Db
643 ValSerAlaArgLysValLeuAlaArgGluGlyGly-----AspGlyAsnProPro 660
Qy
2113 GTGTGTCTAAATATATCTCTCTATCAGAGAGAGAACAGAGAGAACCGGGAACCCACT 2172
Db
661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680
Qy
2173 TTTCAACACTTAGTATCTCTGGCATTAAGCATTAAGTCTTCTCCGAGAGCATCTTCTCAG 2232
Db
681 AspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 700
Qy
2233 GAGAGAAATACCATGGTACCTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 2292
Db
701 AlaAspLysTrpGlyPheGlyValAlaThrThrTrpGluValPheGlnArgGlyProAlaHis 720
Qy
2293 ACAGCAAGTGGAGCTTGGGAGCCACTCTGTGGAGATCTCGAGTGGAGAGATGAAGCCC 2352
Db
721 IleThrSerLeuGluProAlaLysLysLysPheTyrGluAspGlnGlnLeuPro 740
Qy
2353 CTGAGTGTCTGGATTTCTCAAGAGAAAGTCTGAGTCTTATGAAGTATAGCATCAGCTTCT 2412
Db
741 AlaLeuLysThrThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760
Qy
2413 GCACCAAGTGGACAGAGATTAGCAAACTTATAATAATTAATTCATGAGCATATGAGCCAGAT 2472
Db


```

FT conflict replace (1359, G)
FT /*tag= ad
FT conflict replace (1365, Q)
FT /*tag= ae
FT conflict replace (1368, T)
FT /*tag= af
FT conflict replace (1374, T)
FT /*tag= ag
FT conflict replace (1401, C)
FT /*tag= ah
FT conflict replace (1413, C)
FT /*tag= ai
FT conflict replace (1431, T)
FT /*tag= aj
FT conflict replace (1453, Q)
FT /*tag= ak
FT conflict replace (1476, G)
FT /*tag= al
FT conflict replace (1488, T)
FT /*tag= am
FT conflict replace (1511..1512, GT)
FT /*tag= an
FT conflict replace (1578, C)
FT /*tag= ao
FT conflict replace (1590, T)
FT /*tag= ap
FT conflict replace (1593, T)
FT /*tag= aq
FT conflict replace (1602, G)
FT /*tag= ar
FT conflict replace (1623, T)
FT /*tag= as
FT conflict replace (1642, G)
FT /*tag= at
FT conflict replace (1657, C)
FT /*tag= au
FT conflict replace (1728, G)
FT /*tag= av
FT conflict replace (1743, C)
FT /*tag= aw
FT conflict replace (1755, C)
FT /*tag= ax
FT conflict replace (1770, A)
FT /*tag= ay
FT conflict replace (1809, G)
FT /*tag= az
FT conflict replace (1816, G)
FT /*tag= ba
FT conflict replace (1821, C)
FT /*tag= bb
FT conflict replace (1857, A)
FT /*tag= bc
FT conflict replace (1878, T)
FT /*tag= bd
FT conflict replace (1935, A)
FT /*tag= be
FT conflict replace (1938, A)
FT /*tag= bf
FT conflict replace (1963, T)
FT /*tag= bg
FT conflict replace (1974, G)
FT /*tag= bh
FT conflict replace (2025, T)
FT /*tag= bi
FT conflict replace (2055, G)
FT /*tag= bj
FT conflict replace (2079, C)
FT /*tag= bk
FT conflict replace (2082, C)
FT /*tag= bl
FT misc_feature
FT /*tag= bm
FT /note= "Nucleotides corresponding to 7 amino acid insert"

FT conflict replace (2253, A)
FT /*tag= bn
FT conflict replace (2259, G)
FT /*tag= bo
FT conflict replace (2283, A)
FT /*tag= bp
FT conflict replace (2285, C)
FT /*tag= bq
FT conflict replace (2433, G)
FT /*tag= br
FT conflict replace (3453, C)
FT /*tag= bs
FT 3'UTR 3484..3629
FT /*tag= bt
FT conflict replace (3579, C)
FT /*tag= bu
FT conflict replace (3595, A)
FT /*tag= bv
FT conflict replace (3599, A)
FT /*tag= bw
FT conflict replace (3620, A)
FT /*tag= bx

XX US6210654-B1.
PN XX
XX
XX 03-APR-2001.
XX
XX 08-OCT-1997; 97US-0946994.
XX
XX 18-JUN-1996; 96US-0665574.
XX 29-JUL-1993; 93US-0097997.
XX
XX (SJD-) ST JUDE CHILDREN'S HOSPITAL.
XX
XX Ihle J, Witthuhn BA, Quelle FW, Silvennoinen O;
XX WPI; 2001-265367/27.
XX P-PSDB; AAE00352.
XX
XX Modulating a biological response mediated by Jak kinase 2 activation to
XX a cytokine, useful for treating excessive proliferation of eukaryotic
XX cells, comprises inhibiting or enhancing tyrosine kinase activity of
XX Jak kinase in the cell -
XX
XX Example 1; Fig 1; 100pp; English.
XX
XX The present sequence is a cDNA encoding murine (Janus kinase 2) Jak2
XX tyrosine kinase. Jak2 sequence has a 600 amino acid long N-terminus that
XX lacks obvious SH2 (src homology 2) and SH3 domains. Following this is a
XX kinase related domain (domain 2) and a carboxyl kinase domain (domain
XX 1). Jak kinases mediate cytokine activity through their tyrosine

Alignment Scores:
Pred. No.: 1,76e-190 Length: 3629
Score: 2619.50 Matches: 537
Percent Similarity: 65.50% Conservative: 190
Best local Similarity: 48.38% Mismatches: 350
Query Match: 44.70% Indels: 33
DB: 22 Gaps: 15

US-09-397-967A-16 (1-1099) x AAD03607 (1-3629)
QY 4 ProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSerSerGluAla 23
Db 145 CCTGTACATCAGAAATGGTGATATTCCTGGAAGTGCTAATTCCTGTGAAGCAGATAGGCCA 204
QY 24 GlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlyProGlnArgLeuSerPhe 43
Db 205 GTCCTTCAAGTGATCTGTACATTCCTCTGGCAAGCTGAAGGAGAGTATCTGAAGTTT 264
QY 44 SerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAlaCysGlyIle 63
Db 265 CCAAGTGGAGATGTTGTGACAGAGAAATTTGTGGCTGCTCTTAAGCTTGTGGTATT 324

```

Qy	64	LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPhePro	83
Db	325	ACGCCTGTGATCATATAATATGTTTGGCTTAATGAGTGAAACCCGAAAGGATCTGGTACCCA	384
Qy	84	ProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg	103
Db	385	CCCAATCATGTCCTTCACATAGACGAGTCAACAGGATGATGATCTCTACAGGATTAAGG	444
Qy	104	PheTyrPheProAspTyrPhe-----GlyLeuGluThrCysHisArgPheGlyLeuArg	121
Db	445	TTCTACTTCCTCATTTGGTACTGTAGTGGCAGCAGCAGAACTTACAGATACGCGAGTCC	504
Qy	122	LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisIleuPheAlaGlnHis	141
Db	505	CGTGGGGTGAAGCTCCTCTGCTTGATGACTTGTGTCATGTCTTACCTTTTGTCTAGTGG	564
Qy	142	ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGlu	161
Db	565	CGGCATGATTTTGTTCGGGTGATGAGTAAAGTACCTGTGACTCATGAACTCAGGAAGAG	624
Qy	162	PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgPro	181
Db	625	TGTCCTGGGATGGCGGTGTAGACATGATGAGATAGCTTAAGGAGAAACACAGACTCCA	684
Qy	182	GlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal	201
Db	685	CTGGCTGTCTATAACTCTGTCAGTCAAGACATCTTACCAAAGTCGCTTCAGGCGAAG	744
Qy	202	IleGlnGlyGlnAsnPheValThrArgArgArgIleArgThrValValLeuAlaLeu	221
Db	745	ATCCAAGACTATCACATTTTAAACCGGAGCGAATCAGGTACAGATTTCCGACATTCAIT	804
Qy	222	LeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAsp	241
Db	805	CAGCAATTCAGTCAATGTAAAGCCACTGCCAGGAACCTAAACCTTAAGTATCTTAAAC	864
Qy	242	LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal-----	256
Db	865	CTGAAACCCCTGCAGTCGCTTCTACACAGAACAGTTTGAGTAAAGAAATCTGCAAGA	924
Qy	257	GlyLeuProGlyAlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle	276
Db	925	GGT---CCTTCAGGTGAGGAGATTTTTCACCAACCATTAATAAATCTGCAACGGTGAATT	981
Qy	277	ProTrpSerSer-----AsnAspGluLeuPheGlnThr	287
Db	982	CAGTGGTCAAGAGGGAAACATAAGAAAGTGAGACACTGACAGAACAGGACGTACAGTTA	1041
Qy	288	PheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyPro	307
Db	1042	TATTGTGATTTCCCTGATATTATTGATGTAGTATTAAAGCAAGCAACACCG--GAATGC	1098
Qy	308	AlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGlu	327
Db	1099	TCAAAATGAAGTAGAAATTTGTAACTGTCCATAAACAAGATGGTAAGTTTGGAGATAGAA	1158
Qy	328	PheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu	347
Db	1159	CTTAGCTCATTTAAAGAAGCCCTTGTCAATTCGTGTCATTAAATGACGGGTATTACAGACTA	1218
Qy	348	IleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGluGlu	367
Db	1219	ACTGGGGATGGCACCATTAATCTCTGCAAGAGAGTGGCTCCCGAGCTGTGCTCGAGAAC	1278
Qy	368	GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAla	387
Db	1279	ATACACAGCAACTGCCAGCGCCCAATATCAATATGATTTTGCAATTAGCAAACTAAAGAG	1338
Qy	388	AlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPhe	407
Db	1339	CGGGGTAAACGACTGGAATATATGCTACTGATCGACGCCCTTAAGGACTTCAACAAATAC	1398

Qy	408	LeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrIleGlyCysLeuIleArg	427
Db	1399	TTTTCTGACCTTTGCTGTTGGAGGAGAAAATGTCATTCAATATATAACACACTGTTTGGATTACG	1458
Qy	428	GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg	447
Db	1459	AGAATGAGAAATGGAGAAATACAACCTTCAGCGGACCTAAGAGGAACCTTCAGTAACCTTTAAG	1518
Qy	448	GluLeuLeuAlaAlaCysTyrAsnSerGlyLeuArgValAspGlyAlaLeuTyrLeu	467
Db	1519	GACCTTTTGAATTCCTACCATGGAACCTGTCGGCTCACACAGTATCATCTTCCAGTTT	1578
Qy	468	ThrSerCysCysAlaProArgProLysGlyLeuSerAsnLeuIleValValArgArgGly	487
Db	1579	ACCAAAATGCTGCCGCCCAAGCAAGATCAAACTTCCTCGCTCTTCAG-----	1632
Qy	488	CysAsnProAlaProAlaProGlyCysSerProSerCysCysAla-----LeuThr	504
Db	1633	ACAATGGTATTCTGATGTTCCAGATCTCACCAACATTACAGAGGCATAATAATGTGAAT	1692
Qy	505	GlnLeuSerPheHisThrIleProThrAspSerLeuGluThrPheHisGluAsnLeuGlyHis	524
Db	1693	CAAAATGCTGTTTCACAAAATCAGAAATGAAGATTAAATTTAATGAAGACTCTTGCCCAA	1752
Qy	525	GlySerPheThrIlePheArgGlySerArgGluValValAsp----GlyGluThr	543
Db	1753	GGTACTTTTCAAAAATTTTTAAAGGTGAAGAGAGAGAGTTCGAGATTATGTCCTCACTG	1812
Qy	544	HisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGluSer	563
Db	1813	CACAAAACGGAAGTCTTTTGAAGTCTCTAGATAAGCACATAGGAACCTATTTCAGAGTCT	1872
Qy	564	PheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLeuHis	583
Db	1873	TTCTTCGAGCAGCAGCATGATGAGTCAGCTTCTCACAAAGCATTTGGTTTGAATAT	1932
Qy	584	GlyValCysMetAlaGly--AspSerIleMetValGlnGluPheValTyrLeuGlyAla	602
Db	1933	GGTCTCTGTGCTGTGGAGAGGAGAACATCTCTGGTTCAAGAAATTTGTAAAAATTTGGATCA	1992
Qy	603	IleAspMetTyrLeuArgLysArgGlyPheHisLeuValSerAlaSerTyrProLysLeuGluVal	622
Db	1993	CTGATACATACCTTGAAGAGAACAAAAATTCCTATAATATATTATTCGAACTTTGGAGTG	2052
Qy	623	ThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsn	642
Db	2053	GCTAAGCAGTTGGCATGGGCCATGCATTTCTAGAGAAAAATCCCTTTATTCATGGGAAT	2112
Qy	643	ValSerAlaArgLysValLeuLeuAlaArgGluGlyGly-----AspGlyAsnProPro	660
Db	2113	GTGTGTGCTTAAATAATATCTGCTTATCAGAGAGAGAGACAGGAGAACCGGGAACCCACCT	2172
Qy	661	PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr	680
Db	2173	TTTCATCAACCTTAGTGATCTCTGGCATTAGCATTCACGTTCTACCGAAGGACATTTCTTCAG	2232
Qy	681	AspArgIleProTyrValAlaProGlyCysLeuGlnGluAlaGlnThrLeuCysLeuGlu	700
Db	2233	GAGAGAATACCATGGGTACTCTCTGAATGCATTGAGAAATCCTAAAAATCTCAATCTGGCA	2292
Qy	701	AlaAspLysTyrGlyPheGlyAlaThrThrTyrGluValPheGlnArgGlyProAlaHis	720
Db	2293	ACAGACAAGTGGAGCTTCGGGACCACTCTGTGGGAGATCTGCTGAGGAGAGATAAGCCC	2352
Qy	721	IleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeuPro	740
Db	2353	CTGAGTGTCTGGATTCTCAAGAGAAAGCTGCAGTCTCTATGAAGATAAGCATCAGCTTCCT	2412
Qy	741	AlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly	760
Db	2413	GCACCAAGTCGACAGAGTAGCAAAACCTTATAATTAATTCATGGCATATGAGCCAGAT	2472
Qy	761	ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr	780

```
Db 2473 TTTCAGGCGCTGCTTTCAGAGCTGTTCATCCGCTGATCTTAACAGCCTGTTTACTCCAGATTAT 2532
QY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCysVal 799
Db 2533 GAACCTACTAACAGAA---AATGACATGCTACCAAAACATGATAATAGTGCCTCAGGGTTT 2589
QY 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLeu 819
Db 2590 TCTGGTGTCT-----TTTGAAGACAGGACCTTACAGATTGGAAGAGACACTTGAAG 2643
QY 820 TyrIleSerLeuLeuGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIle 839
Db 2644 TTCTACAGCAGCTTGGCAAGGTAATCTCGGAGGTGCGAGATGCGCGCTATGACCCG 2703
QY 840 LeuGlyAspAsnThrGlyProLeuValAlaValIleGlnLeuGlnHisSerValProAsp 859
Db 2704 CTGAGGACAACTCTGGCAGGTGGTGTGTGAAGAACTCCAGCACAGCACTGAAGAG 2763
QY 860 GlnGlnArgAspPheGlnArgGlnIleGlnIleLeuLeuAlaLeuHisSerAspPheIle 879
Db 2764 CACCTCCGAGACTTTGAGAGGAGATCGAGATCTGAATCTTTCAGCATGACACATC 2823
QY 880 ValIleTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuMetGlu 899
Db 2824 GTCAAGTACAGGGAGTGTCTACAGTGGCGGTGGCGCAACCTAAGATTAATATGGAA 2883
QY 900 TyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHisThr 918
Db 2884 TATTTACCATATGGAGTTTACGAGACTATCTCCAAAACATAAAGACGGATAGATCAC 2943
QY 919 AspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArg 938
Db 2944 AAAAAACTTCTTCAATACATCTCAGATATGCAAGGGATGGAATATCTTGGTACAAA 3003
QY 939 ArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisVal 958
Db 3004 AGGTATATCCAGGGACCTGGCAACAGGAACATATTGGTGAAATGAGAACAGGGTT 3063
QY 959 LysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrValVal 978
Db 3064 AAAATAGGAGACTTCGATATACCAAGTCTTGGCCGACGAAAGAACTACTACAAAGTA 3123
QY 979 ArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnIle 998
Db 3124 AAGGAGCCAGGGGAAAGCCCATATTCTGTGTACGACCTGAATCTTTCAGGAGCAG 3183
QY 999 PheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
Db 3184 TTTTCTGTGGCTCAGATGTGTGGAGCTTTGGAGTGGTTCTATACGAATTTTCATAC 3243
QY 1019 CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetClyProGluArgGlu 1038
Db 3244 ATCGAAGAGTAAAGTCCACCGTGGAAATTTATCGGAATGATGGCAATATAAACAA 3303
QY 1039 GlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProPro 1057
Db 3304 GGGCAATGATGTGTTCATTGTAGAGCTACTGAAGAGCAACCGAAGATGCCAAGG 3363
QY 1058 ProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGluPro 1077
Db 3364 CCAGAAGGATGCCAGATGAGATTTATGTGATCATGACAGAGTGTGGGAACAACATGTG 3423
QY 1078 HisAspArgProAlaPheAlaThrLeuSer 1087
Db 3424 AGCCAGCGTCCCTCCTTCAGGACCTTCG 3453

RESULT 12
ABQ76194
ID ABQ76194 standard; DNA; 3503 BP.
XX
AC ABQ76194;
XX
```

```
DT 18-OCT-2002 (first entry)
XX Human Jak2 cDNA.
XX
XX Jak2 kinase; human; immunosuppressive; cytostatic; chromosome 9p23-24;
KW autoimmune disorder; malignancy; leukaemia; gene therapy; vaccine;
KW Jak2; ss.
XX
XX Homo sapiens.
XX
XX CA2203706-A.
XX
XX 25-OCT-1998.
XX
XX 25-APR-1997; 97CA-2203706.
XX
XX 25-APR-1997; 97CA-2203706.
XX
XX (HSCR-) HSC RES & DEV LP.
XX
XX Roifman CM;
XX
XX WPI; 2000-024038/03.
XX
XX P-PSDB; ABB84401.
XX
XX Isolated nucleic acid encoding a human kinase, useful for treating
XX autoimmune disorders and malignancies -
XX
XX Disclosure; Page 44-46; 57pp; English.
XX
XX This invention describes a novel isolated nucleic acid encoding a human
XX Jak2 kinase which has immunosuppressive and cytostatic activity. The Jak2
XX kinase of the invention is located on chromosome 9p23-24 and is expressed
XX highly in spleen, lymph nodes and peripheral blood lymphocytes.
XX Expression of the Jak2 polypeptide increases dramatically on activation
XX of mature B lymphocytes. Jak2 and its mutants are useful for treating
XX autoimmune disorders and malignancies such as leukaemia and can be used
XX for gene therapy or vaccine production. This sequence encodes the human
XX Jak2 protein described in the method of the invention.
XX Note: This specification contained no claims.
XX
XX Sequence 3503 BP; 1180 A; 612 C; 741 G; 970 T; 0 other;

Alignment Scores:
Pred. No.: 4,34e-188 Length: 3503
Score: 2588.00 Matches: 526
Percent Similarity: 66.27% Conservative: 191
Best Local Similarity: 48.61% Mismatches: 335
Query Match: 44.16% Indels: 30
DB: 21 Gaps: 15

US-09-397-967A-16 (1-1099) x ABQ76194 (1-3503)
QY 36 GlyProProGlnArg-----LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeu 53
Db 142 GGGAACTTGAGCAGATTAATCTGACCTTCCATCTGGGAGTATGTTGCAGAAAGAAATC 201
QY 54 CysValArgAlaAlaIleAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeu 73
Db 202 TGTATTTGCTGCTTCAAAGCTTGTGTATCACACCTGTGTATCAATAATATGTTGCTTTA 261
QY 74 AlaThrGluAspPheSerCysTrpPheProSerHisIlePheCysIleGluAspVal 93
Db 262 ATGAGTGAACAGAAAGGATCTGGTATCCACCCACCATGCTTCCATATATAGATGATCA 321
QY 94 AspThrGlnValLeuValTyrArgLeuArgPheTyrPheProAspTrpPhe-----Gly 111
Db 322 ACCAGGCATATGACTCTACAGATAAGATTTTACTTTCCTCGTGTGATTTGCAGTGGC 381
QY 112 LeuGluThrCysHisArgPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeu 131
Db 382 AGCAACAGAGCCTATCGGCATGGAATATCTCGAGGTGCTCGAGCTCCTCTTCTTGTATGAC 441
```

132 HisValLeuGluHisLeuPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuPro 151
152 ValGlyLeuSerMetLysGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAla 171
502 GTACCTGTGACTCGTGAACACAGGAAGAATGCTTGGGAGCGCAGGTGTACATATGATG 561
172 GlnMetAlaArgGluGlnAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLys 191
562 AGAATAGCCAAAGAAACGATCAACCCCACTGGCCATCTATAACTCTATCAGCTACAG 621
192 AlaCysLeuProSerLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArg 211
622 ACAATCTTACCAAAATGTATTTCGAGCAAGATCCAGACTATCATATTTTGACAAGGA 681
212 ArgIleArgThrValValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgPro 231
682 CGAATAAGGTACAGATTTCGCAGATTATTTCAGCAATTTCAGCAATTCAGCAATGCA 741
232 TyrAlaLeuMetAlaLysTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThr 251
742 AGAACTTGAACTTAAGTATCTTATAAATCTGAAACTCTGCAGCTGCGCTCTACACA 801
252 GluThrPheArgValGlyLeuProGlyAla-----GlnGluGluProGlyLeu 267
802 GAGAAATTTGAAGTAAAAAGAACCTTGAAGTGGTCTCTTCAGTGCAGGAGATTTTTCGA 861
268 LeuArgValAlaGlyAspAsnGlyIleProThrSerSer----- 280
862 ATTATAATACTGGAACCGTGGAAATTCAGTGGTCAAGAGGGAACATATAAGAAAGTG 921
281 -----AsnAspGluLeuPheGluThrPheCysAspPheProGluIleValAspValSer 298
922 ACCTGCACAGACGAGATTACAGTTATATTCGATTTTCTATAATATATGATGTCAGT 981
299 IleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArg 318
982 ATTAAGCAAGCAAAACCAAGAGGGT---TCAAAATGAAAGCGGAGTTGTAACATATCCA 1038
319 MetAspGlyHisIleLeuGlu---AlaGluPheProGlyLeuProGluAlaLeuSerPhe 337
1039 CAAGATGTGTAAAAATCTCGAAATATTGAACATTAGCTCATTAAGGGAAGCTTTGTCT 1098
338 ValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLys 357
1099 GTGTCTATTAAATGTATGATATATTAGATTAACTGCAGATGCACATCATTAACCTCTG 1158
358 GluValAlaProArgLeuLeuGluGluAlaAspValCysHisGlyProIleThr 377
1159 GAAGTAGCACCTCCAGCCGTGCTTGAAATATATACAAAGCAACTGTCATGGCCCAATTTC 1218
378 LeuAspPheAlaIleHisLysLeuLysAlaIleGlySerLeuProGlyThrTyrIleLeu 397
1219 ATGTTTGTTCATAGTAAACTGGAAGAACCGAGTAATCAGACTGCAGCTGTATGTACTA 1278
398 ArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeu 417
1279 CGATCGATCCGAGAGACTTTAATAATATTTTCTGACCTTTGCTGTGTCAGCGAGAAAT 1338
418 GlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuVal 437
1339 GTCATTGAATAATAACACTGTTTGTATTACAAAAATGAGATGAAGAGTACAACCTCAGT 1398
438 GlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGly 457
1399 GGGACAAAAGAAACCTTCAGCAGCTCTTAAGATCTTTTGAATCTTTACCAGATGAAACT 1458
458 LeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGlu 477
1459 GTTCGCTCAGACAAATATAATTTTCCAGTTTACTAAATCTCTTAAATCTGTCCCCAACG 1518
478 LysSerAsnLeuIleValValArg--ArgGlyCysAsnProAlaProAlaProGlyCys 496

Db	1519	AAATCAAAACCTTCTAGTCTTCAGAACGAATGGTGTCTCTGATGTACCAACCTCACCACA	1578
Qy	497	SerProSerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeu	516
Db	1579	TTACAGAGGCTACTCATATGAACCAAAATGGTGTTCACAAAATCAGAAATGAAGATTG	1638
Qy	517	GluTrpHisGlnAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArg	536
Db	1639	ATATTTAATGAAGCCCTGGCCAAAGGCACATTTTACAAAGATTTTAAAGCGGTACCGAGA	1698
Qy	537	GluValValAsp--GlyGluThrHisAspSerGluValLeuLeuLeuLysValMetAspSer	555
Db	1699	GAAGTAGGAGACTACGCTCAACTGCATGAACACAGAGTCTCTTTAAAGTTCGGATAAA	1758
Qy	556	ArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSer	575
Db	1759	GCACACAGAAACATTTCAGAGTCTTCTTTGAAGCAGCAAGTATGATGACCAAGCTTCT	1818
Qy	576	TyrProHisLeuValLeuLeuLeuHisGlyValCysMetAlaGlyAsp---SerIleMetVal	594
Db	1819	CACAAGCATTTGGTTTTAAATTTATGGAGTATGTCTGTGGAGACGAGAATATTATGGTT	1878
Qy	595	GlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuVal	614
Db	1879	CAGAGATTGTAAATTTGGATCATCTAGATACATATCTGAAAAAGAAATAAAATTTGTATA	1938
Qy	615	SerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlu	634
Db	1939	AAATATATATGAAACCTTGAAGTGTCTAAACAGTTGGCATGGGCCATTTCTTAGAA	1998
Qy	635	AspLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGly	654
Db	1999	GAATAACACCTTATTCATGGGAATGTATGTGCCAAAAATATTCAGCTTATCAGAGAA	2058
Qy	655	Gly-----AspGlyAsnProPheIleLysLeuSerAspProGlyValSerProThr	672
Db	2059	GACAGGAAGACAGAGAAACCTCTCTTTCATCAACTTAGTGACCTGGCATTAGTATTACA	2118
Qy	673	ValLeuSerLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGln	692
Db	2119	GTTTGTCCAAAGGACATTTCTCAGGAGAGATACCATGGGTACCACCTGAATGCATTGAA	2178
Qy	693	GluAlaGlnThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGlu	712
Db	2179	AATCTTAAAAATTTAAATTTGGCAACACAGACAAATGGAGTTTGGTACAACCTCTGTGGAA	2238
Qy	713	ValPheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPhe	732
Db	2239	ATCTGCAGTGGAGAGATAAACCTCTTAAGTGCTCTGGATTCTCAAGAAAGCTACAAATT	2298
Qy	733	TyrGluAspGlnGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThr	752
Db	2299	TATGAAGATAGCATCAGCTCTTCGCACCAAGTGGGCAGAAATTAGCAACCTTATAAAT	2358
Qy	753	GlnCysMetAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeu	772
Db	2359	AAATTGTATGGATTATGAACACAGATTTACGGCTCTCTTTCAGAGCATCATACAGAGATCTT	2418
Qy	773	AsnGlyLeuIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSer	792
Db	2419	AACAGTTTGTCTTACTCCAGATTATGAATCTATTAACAGAA---ATGACATGTTCACCAAT	2475
Qy	793	ProArgAsp--GluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAla	811
Db	2476	ATGAGAAATGGTGCCTAGGTTTCTGGTCC-----TTTGAACCGGATCCTACA	2522
Qy	812	IlePheGluGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySer	831
Db	2530	CAGTTTGAAGAGAGACATTTGAAATTTCTCAGCAACCTTGGCAAGGGTAAATTTTGGGAGT	2589
Qy	832	ValGluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLys	851

Db 2590 GTGGAGATGTCGGGTATGACCTCTACAGACACACTGGGAGGTGGTGGCTGTAAAA 2649

Qy 852 GlnLeuGlnHisSerValProAspGlnGlnArgAspPheGlnArgGluLeuGlnLeu 871

Db 2650 AGCTTCACATAGTACTGAAGACACCTTAAGAGACTTTGAAAGGAAATTTGAAATCCCTG 2709

Qy 872 LysAlaLeuHisSerAspPheLeuValLysTyrArgGlyValSerTyrGlyProGlyArg 891

Db 2710 AAATCCTACAGCATGACACACATTTAAAGTACAGGAGTGTGTACAGTGTGTCTGG 2769

Qy 892 GlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGln 911

Db 2770 CTAATCTTAAGTTAATTAATGAATATTTACCATATGGAAGTTTACGAGATATCTTCAA 2829

Qy 912 ArgHisArg--GlyLeuHisThrAspArgLeuLeuPheAlaTrpGlnLysCysLys 930

Db 2830 AAACATAAAGACGGATAGATCACATAAACTTCTGCAGTACACACTCTCAGATATCAAG 2889

Qy 931 GlyMetGluTyrLeuGlyAlaArgCysValHisArgAspLeuAlaAlaArgAsnLeu 950

Db 2890 GGTATGGTGTATCTTGGTACAAAAGGTATATCCACAGGATCTGGACAGAGAATATA 2949

Qy 951 LeuValGluSerGluAlaHisValLysLeuAlaAspPheGlyLeuAlaLysLeuPro 970

Db 2950 TTGGTGGAGAACGAGAACAGAGTAAATTTGGGTTAACCAAGTCTTGCCA 3009

Qy 971 LeuGlyLysAspTyrValValArgGluProGlyGlnSerProLeuPheTrpValAla 990

Db 3010 CAAGCAAGAATATCTTAAGTAAAGAACCTGGTAAAGTCCCATATCTGGTATGCT 3069

Qy 991 ProGluSerLeuSerAspAsnLeuPheSerArgGlnSerAspValTrpSerPheGlyVal 1010

Db 3070 CCAGAATCACTCACAGAGAGCAAGTTTCTGTGGCCTCAGATGTTTGGAGCTTGGAGTG 3129

Qy 1011 ValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeu 1030

Db 3130 GTTCTGTATGAATCTTTTCATACATCTGAGAGAGTAAAGTCCACAGCGGAATTTATG 3189

Qy 1031 ArgMetMetGlyProGluArgGluGlyProProLeuCys--ArgLeuLeuGluLeuLeu 1049

Db 3190 CGTATGATGGCAATGACAAACAGGACAGATGATGCTGCTTCATTTGATGAACCTTTG 3249

Qy 1050 AlaGluGlyArgArgLeuProProProThrCysProThrGluValGlnGluLeuMet 1069

Db 3250 AAGAATAATGGAAGATTACCAAGACAGATGGATGCCAGATGAGATCTATATGATCATG 3309

Qy 1070 GlnLeuCysTrpAlaProGluProHisAspArgProAlaPheAlaThrLeuSerProGln 1089

Db 3310 ACAGATGCTGGAAACAATAATGTAATCAACGCCCTCTTGGGATCTAGTCTCTCGA 3369

Qy 1090 LeuAsp 1091

Db 3370 GTGGAT 3375

RESULT 13

ACA55699

ID ACA55699 standard; cDNA; 3900 BP.

XX AC ACA55699;

AC ACA55699;

XX 06-JUN-2003 (first entry)

DT Human signalling pathway polynucleotide probe SEQ ID NO 297.

DE Human; probe; ss; array element; Parkinson's disease;

XX signalling pathway population; cancer; adenocarcinoma; leukaemia;

KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.

XX Homo sapiens.

OS US6500938-B1.

PN 31-DEC-2002.

XX PD

XX 30-JAN-1998; 98US-0016434.

XX 30-JAN-1998; 98US-0016434.

XX (INCY-) INCYTE GENOMICS INC.

XX Au-Young J, Seilhamer JJ;

XX WPI; 2003-352189/33.

XX Combination of polynucleotide probes, useful as array elements in a

PT microarray for monitoring the expression of a number of target

PT polynucleotides -

XX Claim 1; SEQ ID NO 297; 65pp; English.

XX The invention relates to a combination which, comprises a number of

CC polynucleotide probes comprising a sequence selected from one of the 1490

CC sequences mentioned in the specification. The combination is useful as an

CC array element in a microarray for monitoring the expression of a number

CC of target polynucleotides. The microarray is particularly useful in the

CC diagnosis and treatment of cancer and immunopathology and neuropathology.

CC The microarray is useful in diagnostics and treatment regimens, drug

CC discovery and development, toxicological and carcinogenicity studies,

CC forensics and pharmacogenomics. The microarray is also useful for

CC monitoring progression of diseases and for developing sophisticated

CC profiles for the effects of currently available therapeutic drugs. The

CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs

CC and genomic fragments and in research and diagnostic applications. The

CC array can detect changes in expression in a large number of genes coding

CC for different signaling pathway populations which can be used to diagnose

CC various diseases including cancer e.g. adenocarcinoma and leukaemia,

CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease

CC and Parkinson's disease. The present sequence represents a polynucleotide

CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification but was obtained in electronic format directly from USPTO

CC at seqdata.uspto.gov/sequence.html?DocID=06500938B1.

XX

SEQ Sequence 3900 BP; 1267 A; 714 C; 853 G; 1048 T; 18 other;

Alignment Scores:

Pred. No.: 2,968-180 Length: 3900

Score: 2486.50 Matches: 509

Percent Similarity: 64.83% Conservative: 184

Best Local Similarity: 47.61% Mismatches: 329

Query Match: 42.43% Indels: 48

DB: 25 Gaps: 14

US-09-397-967A-16 (1-1099) x ACA55699 (1-3900)

Qy 48 TyrLeuAlaGluAspLeuCysValArgAlaAlaLysAlaCysGlyIleLeuProValTyr 67

Db 463 TATCTTACCATTCCCTTTGATTGCTGCTTCAAGCTTGGTATCACACTGTAT 522

Qy 68 HisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPheProProSerHisIle 87

Db 523 CATAATATGTTGCTTTTAAATGAGTCAAAACAGAAAGGATCTGTTATCCACCAACCATGTC 582

Qy 88 PheCysIleGluAspValAspThrGlnValLeuValTyrArgLeuArgPheTrpPhePro 107

Db 593 TTCCATATAGTAGTCAACCCAGGCAATATGATCTCTACAGAAATGATTTTACTTTCT 642

Qy 108 AspTrpPhe-----GlyLeuGluThrCysHisArgPheGlyLeuArgLysAspLeuThr 125

Db 643 CG-TGTTATTCAGCTGGGAGCAACAGACCTATCGGATGGAATATCTCGAGGTCTCAA 701

Qy 126 SerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHisArgSerAspLeu 145

Db 702 GCTCCTCTTCTGTAGTACTTTGTCATGCTTACCTTTTGTCTCAGTGGCGCATGATTTT 761

Qy 146 ValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGluPheLeuSerLeu 165

```
Db 762 GTGCATGGATGATAAAGTACCTGTGACTCATGAACACAGGAAGAAATGCTCTGGGATG 821
Qy 166 AlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgProGlyGluLeu 195
Db 822 ACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 891
Qy 186 LysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspValIleGlnGly 205
Db 882 AACTCTATCAGCTACAGACATCTTACCAACATGTTATCGCAAGATCCCAAGACTAT 941
Qy 206 AsnPheValThrArgArgGileAlaArgThrValValLeuAlaLeuLeuProCysGly 225
Db 942 CATATTTCAGACAGGAAGCGAATTAAGGTACAGATTTTCGCAGATTTAATTCAGCAATTCAGC 1001
Qy 226 ArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGluArgLeu 245
Db 1002 CAATGCAAAACCCACTGCGAGAACTTTGAAACTTTAAGTATCTTATAAATCTGGAACACTCTG 1061
Qy 246 HisProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAla----- 261
Db 1062 CAGTCTGCTTCTACACAGAAATTTGAAGTAAAGAACTCGAAGTGGTCTTCAAGT 1121
Qy 262 GlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrSerSer--- 280
Db 1122 GAGGAGATTTTGCACCACTTATAAATCACTGGAACCGTGAATTCAGTGGTCAAGAGGG 1181
Qy 281 -----AsnAspGluLeuPheGlnThrPheCysAspPhePro 292
Db 1182 AAACATAAAGAAAGTGAGACACTGACAGAACAGGATTTACAGTTATATTCGGATTTTTCCT 1241
Qy 293 GluIleValAspValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArg 312
Db 1242 AATATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1298
Qy 313 LeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuPro 332
Db 1299 GTTGTAATCATCCATAAGCAGATGGTAAATAATCTGGAATTTGAATAGCTCATTAAGG 1358
Qy 333 GluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArg 352
Db 1359 GAAGCTTTGCTCTTTCGTGTCATTAATTTGATGATGATGATGATGATGATGATGATGAT 1418
Qy 353 HisTyrPheCysLysGluValAlaProProArgLeuGluGluAlaAspValCys 372
Db 1419 CATTTACCTCTGTAAGAAGTAGACCTCCAGCCGTGCTTGAATAATATACAAAGCAACTGT 1478
Qy 373 HisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuPro 392
Db 1479 CATGCCCAATTTTCATGATTTTGGCATTAGTAACTGAAGAAAGCAGGTAATCAGACT 1538
Qy 393 GlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCys 412
Db 1539 GGACTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1598
Qy 413 ValGlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGly 432
Db 1599 GTCAGCGCAGAAATGTCATTGAATATTAACACTGTTTGTATTACAAAAATGAGATGAA 1658
Qy 433 AlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAla 452
Db 1659 GAGTACAACCTCAGTGGGACAAAGAAAGAACTTCAGCAGCTTTAAGATCTTTTGAATTTGT 1718
Qy 453 CysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysAla 472
Db 1719 TACAGATGGAACCTGTCGTCTCAGACCAATATAATTTCCAGTTTACTAATGCTGCTCC 1778
Qy 473 ProArgProLysGluLysSerAsnLeuIleValValArg---ArgGlyCysAsnProAla 491
Db 1779 CCAAGGCCAAAAGATAAATCAACCTTCTAGTCTTCAGAACGAATGGTGTCTTCATGATGA 1838
Qy 492 ProAlaProGlyCysSerProSerCysAlaLeuThrGlnLeuSerPheHisThrIle 511
```


Db 418 GAGATTGTTGCAACCACTTATAATACTGGAACCGTGGAAATTCAGTGGTCAAGAGGGAAA 477
QY 281 -----AsnAspGluLeuPheGlnThrPheCysAspPheProGlu 293
Db 478 CATAAGGAAGTGGAGACCTGACAGAACAGAGCGTACAGTATATTTGGTGATTTCCCTGGAT 537
QY 294 IleValAspValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeu 313
Db 538 ATTATTGATGTCAGTATTAGCAAGCAACACAG--GAATGCTCAATGAAGTAGAATT 594
QY 314 ValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGlu 333
Db 595 GTAACGTGTCATAAACAAGATGTTAAAGATTTTGGAGATAGAACTTAGCTCATTAATAAGAA 654
QY 334 AlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHis 353
Db 655 GCCTTGTCATTCGTGTCATTAATTTAGCGGTATTACAGACTTAACCTGGGATGGCCACAT 714
QY 354 TyrPheCysLysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHis 373
Db 715 TACCTCTGCAAGAGGTGGTCCCCAGCTGTGCTCGAGAACAATACACACCAACTGCCAC 774
QY 374 GlyProIleThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGly 393
Db 775 GGCCCAATATCAATGGATTTTGCATTTAGCAAACTAAAGAGGCGGTAAACCGACTGGA 834
QY 394 ThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysVal 413
Db 835 CTATATGTGTACATCGACCCCTAAGCACTTACAAATACCTTTCAGCTTTGCTGTT 894
QY 414 GlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAla 433
Db 895 GAGCGAATAATGTCATTTGAATAATAACACTGTTTGTATTACGAAGATGAGATGGAGAA 954
QY 434 PheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaCys 453
Db 955 TACAACCTCGCGGACTTAATAGAACTTCAGTAACTTAAGGACCTTTTGAATTTGCTAC 1014
QY 454 TrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaPro 473
Db 1015 CAGATGGAACCTGTCGCTCGACAGCATCATCTCCAGTTTCCAAATGCTGCCCCCA 1074
QY 474 ArgProLysGlyLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAla 493
Db 1075 AAGCAAAAGATAAATAACCTTCTGCTCTCAGA-----ACAATGGTATTTCTGAT 1128
QY 494 ProGlyCysSerProSerCysCysAla-----LeuThrGlnLeuSerPheHisThr 510
Db 1129 GTTCAGATCTCACCAACATTACAGAGCATATAAATGTGAATCAATGGTGTTCACAAA 1188
QY 511 IleProThrAspSerLeuGluTyrHisGluAsnLeuGlyHisGlySerPheThrLysIle 530
Db 1189 ATCAGGAATGAAGATTTAATATTAAATGAAGTCTTGCCCAAGTACTTTTACAAAAT 1248
QY 531 PheArgGlySerArgArgGluValValAsp---GlyGluThrHisAspSerGluValLeu 549
Db 1249 TTTAAAGGTGTAAAGAGAGAGTGGAGATTATGTCACATGCGACAAAACCGAAGTCTT 1308
QY 550 LeuLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSer 569
Db 1309 TTGAAGTCTCAGTAAAGACACATAGGAATTTTTCAGAGTCTTTTTCGAAGCAGCAAGC 1368
QY 570 LeuMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGly 589
Db 1369 ATGATGAGTCAGCTTTCTCACAGCAATTTGGTTTGAATATGTTGTTCTGTGCTGGA 1428
QY 590 ---AspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArg 608
Db 1429 GAGGAGAACAATCTCTGTTTCAAGATTTGTAATAATTTTGGATCACTGGATACATACCTGAAG 1488
QY 609 LysArgGlyHisLeuValSerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyr 628
Db 1489 AAGAACAAAAATTCATAATATATATTATGAAACTTTGGAGTGGCTAAGCAGTGGCATGG 1548

QY 629 AlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysVal 648
Db 1549 GCCATGCATTTCTAGAGAAAATCCCTTATTTCATGGGAATGTGTGTCTAAAAATATC 1608
QY 649 LeuLeuAlaArgGluGlyGly-----AspGlyAsnProProPheIleLysLeuSerAsp 666
Db 1609 CTGCTTATCAGAAAGAGACAGAGAACCGGGAACCCACTTTCATCAAACTTAGTAT 1668
QY 667 ProGlyValSerProThrValLeuSerLeuGluMetLeuThrAspArgIleProTrpVal 686
Db 1669 CTTGGCATATGACATTCACAGTCTTACCAGAGCATCTTCTCAGAGAGAAATACCATGGGA 1728
QY 687 AlaProGlyCysLeuGlnGluAlaGlnThrLeuCysLeuGluAlaAspLysTrpGlyPhe 706
Db 1729 CCTCTCTGAATGCAATGAGAATCTTAAATCTCAATCTGGCAACAGCAAGTGGAGTTC 1788
QY 707 GlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHisIleThrSerLeuGluPro 726
Db 1789 GGGACCACTCTGTGGAGATCTGCAGTGGAGAGATAAGCCCTGTAGTCTCTGGATTC 1848
QY 727 AlaLysLysLeuLysPheTyrGluAspGlnGlnLeuProAlaLeuLysTrpThrGlu 746
Db 1849 CAAGAAGCTGAGTCTTATGAAGATAAGCATCAGCTTCTGACCACCAAGTGGACAG 1908
QY 747 LeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGlyArgArgProSerPheArg 766
Db 1909 TTAGCAACCTTATAATAATTCATGGACTATGAGCCAGATTCAGGCCCTGCTTTCAGA 1968
QY 767 AlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyrGluLeuLeuSerAspPro 786
Db 1969 GCTGTATCCTGCTGATCTTAACAGCTGTTTACTCCAGATTTATGAATTAACAGAA-- 2025
QY 787 ThrProGlyIleProSerProArg---AspGluLeuCysValAlaGlyAlaGlnLeuTyr 805
Db 2026 AATGACATGCTACCAACATGAGAAATAGTGGCCCTAGGGTTTCTGGTGTCT-----TTT 2079
QY 806 AlaCysGlnAspProAlaIlePheGluGluArgHisLeuLysTyrIleSerLeuLeuGly 825
Db 2080 GAAGACAGGAGCCCTACACACTTTGAAGAGACACTTTGAAGTTTCTACAGCAGCTTGGC 2139
QY 826 LysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeuGlyAspAsnThrGly 845
Db 2140 AAAAGTAACTTCGGAGTGTGGAGATGTGCGCTATGACCGCTGACAGGCAACACATGGC 2199
QY 846 ProLeuValAlaValLysGlnLeuGlnHisSerValProAspGlnGlnArgAspPheGln 865
Db 2200 GAGGTGCTGCTGTGAAGAACTCCAGCAGCAGCTGAAGAGCACCTCCGAGACTTTGAG 2259
QY 866 ArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleValLysTyrArgGlyVal 885
Db 2260 AGGAGATCGAGATCCTGAATCCTTGCAGCATGACAACTCGTCAAGTACAGGAGTG 2319
QY 886 SerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCys 905
Db 2320 TGCTACAGTTCGGGTGCGCGCAACCTAAAGATTAATATGGAATATTTTACCATATGAAGT 2379
QY 906 LeuArgAspLeuLeuGlnArgHisArg--GlyLeuHisThrAspArgLeuLeuPhe 924
Db 2380 TTACGAGACTATCTCCAAAAACATAAGAACCGGATAGATCACAAAAAACTTCTTCAATAC 2439
QY 925 AlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAsp 944
Db 2440 ACATCTCAGATATGCAAGGCGCATGGAATATCTTGTGTACAAAAAGGTATATCCACAGGAC 2499
QY 945 LeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisValLysIleAlaAspPheGly 964
Db 2500 CTGGCAACAAGGAACATATTTGGTGGAAATGAGAACAGGGTGTAAATATAGAGACTTCGGA 2559
QY 965 LeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArgGluProGlyGlnSer 984
Db 2560 TTAACCAAGTCTTGCGCGAGGACAAAGATATCTCAAAAGTAAAGGAGCCAGGGAAGC 2619

QY 985 ProfilePheTrrpTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAsp 1004
 DB 2620 CCCATATTCGTAGCGCACTCAATCCTTGACGGAGACCAAGTTTCTGTGGCCTCAGAT 2679
 QY 1005 ValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSer 1024
 DB 2680 GTGTGGAGCTTTGGAGTGGTCTTATACGAACCTTTTTCACATACATCGAAGAGTAAAGT 2739
 QY 1025 ProSerAlaGluPheLeuArgMetMetGlyProGluArgGluGlyProProLeuCys--- 1043
 DB 2740 CCACCCGTGGAAATTATCGGAATGATTGGCAATGATAACCAAGGCAAAATGATTGTGTC 2799
 QY 1044 ArgLeuLeuGluLeuAlaGluGlyArgArgLeuProProProThrCysProThr 1063
 DB 2800 CATTTGATAGACTACTGAAGCAACGGAAGATTGCCAAGGCCAAGAGGATGCCAGAT 2859
 QY 1064 GluValGlnGluLeuMetGluLeuCysTrpAlaProGluProHisAspArgProAlaPhe 1083
 DB 2860 GAGATTATATGTCATCATCAGACAGTGTCTGGAACAACAATGTGAGCCGCTCCCTCTTC 2919
 QY 1084 AlaThrLeuSer 1087
 DB 2920 AGGACCTTTTCG 2931
 RESULT 15
 ID AAD22680
 XX AAD22680 standard; cDNA; 3495 BP.
 AC AAD22680;
 XX
 XX 26-FEB-2002 (first entry)
 DE Murine JAK2 tyrosine kinase cDNA.
 XX
 KW Murine; Src homology 2-Bbeta; SH2-Bbeta; neuroprotective; gene therapy;
 KW cell differentiation; nerve regeneration; angiogenesis; embryogenesis;
 KW cytosolic; antitense therapy; drug screening; cellular expression;
 KW immunological disease; neurological disease; apoptosis; diabetes; cancer;
 KW arthritis; JAK2 tyrosine kinase; ss.
 XX
 OS Mus sp.
 XX
 FH Key
 FT CDS
 FT Location/Qualifiers
 FT 1..2982
 FT /*tag= a
 FT /product= "Murine JAK2 tyrosine kinase protein"
 XX
 PN USG312941-B1.
 XX
 PD 06-NOV-2001.
 XX
 PF 26-NOV-1997; 97US-0980080.
 XX
 PR 26-NOV-1997; 97US-0980080.
 XX
 PA (UNMI) UNIV MICHIGAN.
 XX
 PI Carter-Su C, Rui L, Karow DS;
 XX
 XX WPI; 2002-024907/03.
 DR P-PSDB; AAB13587.
 XX
 XX New isolated nucleic acid molecule encoding full length Src homology 2
 PT (SH2)-Bbeta protein, useful for modulating cellular expression of
 PT active SH2-Bbeta protein -
 XX
 PS Example 1; Fig 3A; 71pp; English.
 XX
 CC The invention relates to human Src homology 2 (SH2)-Bbeta protein and its
 CC DNA molecule. SH2-Bbeta is useful in drug screening assays designed to
 CC identify drugs that interfere with the specific binding of JAK2 tyrosine
 CC kinase which is a member of Janus family of tyrosine kinase-interacting
 CC signalling molecule. SH2-Bbeta protein is useful for modulating,

CC preferably reducing cellular expression or intracellular concentration or
 CC availability of active SH2-Bbeta. SH2-Bbeta is useful for treating a
 CC disease associated with undesirable cell growth, differentiation, and
 CC growth factor/cytokine, preferably interleukin, more preferably growth
 CC hormone, platelet derived growth factor, nerve growth factor, epidermal
 CC growth factor responsiveness, and for treating diseases associated with
 CC cell movement. SH2-Bbeta molecule has important implications in cancer
 CC metastasis, nerve regeneration, angiogenesis and embryogenesis and is
 CC useful for preventing apoptosis, and treating diabetes, cancer,
 CC arthritis, immunological diseases, neurological diseases, etc.
 CC The invention also relates to compositions and methods for identifying
 CC cytokine, hormone and growth factor signalling pathway agonists and
 CC antagonists. Human SH2-Bbeta DNA is useful in detecting the presence of
 CC SH2-Bbeta genes in gene transcripts, for detecting or amplifying DNA's
 CC with substantial sequence similarity with SH2-Bbeta homologues and
 CC structural analogues and for gene therapy applications. The present
 CC sequence is murine JAK2 tyrosine kinase cDNA related to the invention.
 XX
 SQ Sequence 3495 BP; 1121 A; 688 C; 794 G; 892 T; 0 other;
 Alignment Scores:
 Score: 1.24e-170 Length: 3495
 Pred. No.: 2360.00 Matches: 489
 Percent Similarity: 65.89% Conservative: 164
 Best Local Similarity: 49.34% Mismatches: 300
 Query Match: 40.27% Indels: 38
 DB: 24 Gaps: 15
 US-09-397-967A-16 (1-1099) x AAD22680 (1-3495)
 QY 128 IleLeuAspLeuHisValLeuGluHisPheAlaGlnHisArgSerAspLeuValSer 147
 DB 1 CTGCTTGATGACTTGTGTCATGCTTACCTTCCCTCAGTGGCGCATGTTTGTTCAC 60
 QY 148 GlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGluPheLeuSerLeuAlaVal 167
 DB 61 GGATGATAAAAGTACCTGTGACTCATGAACTCAGGAAGAGTGTCTTGGGATGCGGTG 120
 QY 168 LeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgProGlyGluLeuLeuLysThr 187
 DB 121 TTAGACATGATGAGAATAGCTAAGGAGAAAGACCAGACTCCACCTGGCTGTCTTAACTCT 180
 QY 188 ValSerTyrLysAlaCysLeuProProSerLeuArgAspValIleGlnGlyGlnAsnPhe 207
 DB 191 GTCAGTACAGACATCTTACCAAGTCGTCGCGCAAGATCCAGACATCATTACATT 240
 QY 208 ValThrArgArgArgIleArgArgThrValValLeuAlaLeuLeuProCysGlyArgLeu 227
 DB 241 TTAACCCGGAAGCAATCAGGTACAGATTTCGACAGATTCATTTCAGCAATTCAGTCAATGT 300
 QY 228 ProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGluArgLeuHisPro 247
 DB 301 AAAGCCACTGCCAGAACCTAAACCTTAAGTATCTTATAAACCTCGAAACCCCTGAGTCT 360
 QY 248 AlaAlaThrThrGluThrPheArgVal-----GlyLeuProGlyAlaGln 262
 DB 361 GCCTTCTACACAGACAGATTGGAAGTAAAGAAATCTGCAAGAGT---CCTTCAGTGTAG 417
 QY 263 GluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrpSerSer----- 280
 DB 418 GAGATTTTTCGAACCATTAATAACTGGAACCGGTGGAATTCAGTGTGTCGAAGGGAAA 477
 QY 281 -----AsuAspGluLeuPheGlnThrPheCysAspPheProGlu 293
 DB 478 CATAGGAAGTGAACACTGACAGACAGACAGACCTGACAGTTATATTGTGATTTCCCTCAT 537
 QY 294 IleValAspValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeu 313
 DB 538 ATTATGATGTCAGTATTAAACAGCAAAATCAG---GAATGCTCACTCAAGATGAGATT 594
 QY 314 ValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGlu 333
 DB 595 GTGACCGTCCCAAGCAGGACGGGAAGGTCTTTGAAATAGAACTTAGCTCTATTAAAGAA 654

Qy	334	AlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHis	353
Db	655	GCCTTGTCATTGCGTGTCAATTAATGACGGGTATTACAGACTTAACCTGCGGATGCACCAT	714
Qy	354	TyrPheCysIysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHis	373
Db	715	TACCTCTGCAAGAGGTGGCTCCCGCCAGCTGTGTTCCGAGAACATACACAGCAACTGCCAC	774
Qy	374	GlyProIleThrLeuAspPheAlaIleHisIysLeuLysAlaAlaGlySerLeuProGly	393
Db	775	GGCCCAATTTCAATGGATTTTGCCATCAGCAAACTAAAGAAGGAGCAAAACACGACCTGGA	834
Qy	394	ThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeuThrAlaCysVal	413
Db	835	CTGTATGTACTTCGATGTAGCCCTAAGGACTTCAACAAATACTTCCTTGACCTTGCGGTT	894
Qy	414	GlnThrProLeuGlyProAspTyrLysGlyCysIleuIleArgGlnAspProSerGlyAla	433
Db	895	GAGCGAGAAATGTTATTGAATATAACACTGTTTGATTAACAAGAATAGAAATGGAGAG	954
Qy	434	PheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCys	453
Db	955	TACAACCTCAGTGGACTAAGAGGAACTTCAGTAGTCTTAAAGGACCTTTTGAATTGCTAC	1014
Qy	454	TrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaPro	473
Db	1015	CAGATGAAACTGTGCGCTCAGACAGTAGTATCATCTTCAGTTCCAAATGCTGCTCTCCA	1074
Qy	474	ArgProLysGluLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAla	493
Db	1075	ANGCGGAAAGATAATCAACCTTCTTGTCTCAGA-----ACAAATGGTGTTCGTGAT	1128
Qy	494	ProGlyCysSerProSerCysCysAla-----LeuThrGlnLeuSerPheHisThr	510
Db	1129	GTTCAGCTCTCACCAACATACAGAGCGCAATAATGTGAATCAATGGTGTTCACAAA	1188
Qy	511	IleProThrAspSerLeuGluTyrPheIleGlnLeuGlyHisGlySerPheThrLysIle	530
Db	1189	ATCAGGAATGAAGATTGTATTTAATGAAGCCTTCGGCAAGCACTTTTACAAAATA	1248
Qy	531	PheArgGlySerArgGluValValAsp---GlyGluThrHisAspSerGluValLeu	549
Db	1249	TTTAAAGGTGTAAAGAGAAATTGGAGATTATGGTCAGCTGCAGCAAAACCGAATTCCT	1308
Qy	550	LeuLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSer	569
Db	1309	TTGAAAGTCCTAGATAAAGCACATAGAAACTATTTCAGAGTCTTTCTTTGAAGCAGCAAGC	1368
Qy	570	LeuMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGly	589
Db	1369	ATGATGAGTCAGCTTCTCACAAGCATTTGGTTTTTGAATTATGAGGATGTGTCTGTGGA	1428
Qy	590	---AspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArg	608
Db	1429	GAGGAGACAATTTTGGTTCAGAGTTTGTAAATTTGGATCAGCTGGATACATACCTGAAG	1488
Qy	609	LysArgGlyHisLeuValSerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyr	628
Db	1489	AAGAACAAAAATCTATAAATATATTATGAAACTTGGAGTGGCGAAGCAGTGTGGCATGG	1548
Qy	629	AlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysVal	648
Db	1549	GCCATGCACTTCTCGAAGAAAAATCCCTATTATCATGGGAATGTGTGCTCTAAAAATATC	1608
Qy	649	LeuLeuAlaArgGluGlyGly-----AspGlyAsnProProPheIleLysLeuSerAsp	666
Db	1609	CTGCTTATCAGAGAAGAAGACAGGAGAACGGGGAACCCACTTTCATCAAACTTAGTGAT	1668
Qy	667	ProGlyValSerProThrValLeu-----SerLeuGluMetLeu	679
Db	1669	CCTGGCATTAGCAATTACAGTTCTTACCGAAGGACATTTCTTCCTGTGTGTTTCCAAATCTT	1728

QY	680	ThrAspArgIleProTrpTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeu	699
DB	1729	CAGGAGAGAATACCATGGGTACCACCTGTAGTCGATTCAGAAATCCTTAAATCATCACTCTG	1788
QY	700	GluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAla	719
DB	1789	GCAACACACAAAGTGGAGCTTCGGGACACACTCTGTGGGAGATCTGCAGTGGAGGAGATAAG	1848
QY	720	HisIleThrSerLeuGluProAlaLysLeuLysPheTyrGluAspGlnGlyGlnLeu	739
DB	1849	CCCCTGAGTGTCTGGATTCTCAAGAAAGCTGCAGTTCATGAGATAAGCATCAGCTT	1908
QY	740	ProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPro	759
DB	1909	CCTGCACCCCAAGTGCACAGAGTTGGCAACCTTATAAATAATTCATGGACTATGAGCCA	1968
QY	760	GlyArgGProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAsp	779
DB	1969	GATTTCAGGCGCTCTTCAGAGCTGTCACTCCGTGATCTTAACAGCTGTTCATCTCCAGAT	2028
QY	780	TyrGluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCys	798
DB	2029	TATGAACACTACTACAGAA---AATGACATCTCTACCAACATGAGATAGTGGCCCTAGGG	2085
QY	799	ValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeu	818
DB	2086	TTTTCTCGTGCT-----TTTGAAGACAGGACCCCTACACAGTTTGAAGAGACACACTTG	2139
QY	819	LysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAsp	838
DB	2140	AAGTTTCTACAGCAGCTTGGCAAGGTAACTCCGGAGTGTGGAGATGTGCCCTATGAC	2199
QY	839	ProLeuGlyAspAsnThrGlyProLeuValAlaValIleGlnLeuGlnHisSerValPro	858
DB	2200	CCGCTGCAGACAACTGCGGAGTGTGCTGTGAAGAACTCCGACACACACTGAA	2259
QY	859	AspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPhe	878
DB	2260	GAGCACCTCCGAGACTTTTGAGGGGAGATCGAGATCCTGAAATCCTTGACGATGACAAC	2319
QY	879	IleValIysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMet	898
DB	2320	ATCGTCAAGTACAAGGGAGTGTCTCAGTGGCGGTGGCGCAACCTTAAGATTAAATATG	2379
QY	899	GluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHis	917
DB	2380	GAATATTTCCATATGGAAGTTTACGAGACTATCTCCAAAACATAGAAGACGGATAGAT	2439
QY	918	ThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAla	937
DB	2440	CACAAAAAATCTTCTCAATACACATCTCAGATATGCAAGGCAATGGAATATCTTGTACA	2499
QY	938	ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis	957
DB	2500	AAAAGGTATATCCACAGGGACCTGGCAACAAGGAACATATTGGTGGAAAATGAGAACAG	2559
QY	958	ValIlyIleAlaAspPheGlyLeuAlaLysLeuLeuLeuProLeuGlyLysAspTyrTyrVal	977
DB	2560	GTTAAATATAGAGACTTCGGATTAAACAAAGTCTTCGGCAGCAAGAAATACTACAAA	2619
QY	978	ValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLeuSerAspAsn	997
DB	2620	GTAAGAGCCAGGGGAACCCCATATCTGTCACGACCTGAATCCITTCGGGAGAGC	2679
QY	998	IlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThr	1017
DB	2680	AAGTTTCTGTGGCCTCAGATGTGTGGAGCTTTGGAGTGTGTTATACGAACCTTTTACA	2739
QY	1018	TyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArg	1037
DB	2740	TACATCCAGAGAGTAAAGTCCACCCGTGGAAATTTATGCGAATGATTTGGCAATGATAA	2799
QY	1038	GluGlyProProLeuCys---ArgLeuLeuLeuAlaGluGlyArgArgLeuPro	1056

```

Db      2800 CAAAGGCAATGATTGTGTTCATTGATAGAGCTACTGAAGAGCAACGGAAGATTGCCA 2859
Qy      1057 ProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGlu 1076
Db      2860 AGGCCAGGAAGGATGCCCGAGATGAGATTATGTGATCATGACAGAGTGCTGTGGAACAACAAT 2919
Qy      1077 ProHisAspArgProAlaPheAlaThrLeuSer 1087
Db      2920 GTGAGCCAGCGTCCCTCCTTCAGGGACCTTCC 2952

```

Search completed: February 4, 2004, 03:08:17
 Job time : 917.095 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2004, 02:45:11 ; Search time 6446.54 Seconds
(without alignments)
4143.406 Million cell updates/sec

Title: US-09-397-967a-16
Perfect score: 5860
Sequence: 1 MAPSEETPLIPQRSCLSS.....RPAFATLSQLDPLWRGPG 1099

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09397967/runat_03022004_175632_28586/app_query.fasta_1.1486
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967@cgn_1_1_4514@runat_03022004_175632_28586 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pbg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	5669.5	96.7	3742	11	AK038268	AK038268 Mus muscu
2	5650	96.4	3781	11	AK088365	AK088365 Mus muscu
3	1447	24.7	4462	11	BC053021	BC053021 Mus muscu
4	1380	23.5	971	12	BT411962	BT411962 602966375
5	1358	23.2	1625	11	BC027234	BC027234 Mus muscu
6	1214.5	20.7	1200	13	BX415899	BX415899 BX415899
7	1192	20.3	826	12	B1557690	B1557690 603237050
8	1170.5	20.0	2222	11	AK087255	AK087255 Mus muscu
9	1141	19.5	698	12	BM935182	BM935182 UI-M-BH3-
10	1134	19.4	675	10	BF150250	BF150250 uy83e09.y
11	1081	18.4	677	14	BY730320	BY730320 BY730320
12	1071.5	18.3	927	13	BX370312	BX370312 BX370312
13	1032	17.6	655	14	BY746452	BY746452 BY746452
14	1031.5	17.6	907	12	BG870513	BG870513 602791432
15	1028.5	17.6	720	12	BG873355	BG873355 602794328
16	1016	17.3	621	9	AW742526	AW742526 up58a12.y
17	1012	17.3	731	9	AA920299	AA920299 vx93902.r
18	1007	17.2	666	14	BY734043	BY734043 BY734043
19	960	16.4	583	10	BB631706	BB631706 BB631706
20	935.5	16.0	871	13	B0709825	B0709825 AGENCOURT
21	930	15.9	559	12	BG800853	BG800853 0053-91 M
22	906	15.5	769	9	AJ442008	AJ442008 AJ442008
23	879	15.0	545	10	BE916148	BE916148 601666763
24	874	14.9	539	10	BE913247	BE913247 601668246
25	835.5	14.3	543	10	BE753801	BE753801 206963 MA
26	831.5	14.2	562	10	BF193178	BF193178 244670 MA
27	828	14.1	580	10	B9608771	B9608771 BB608771
28	824	14.1	557	12	B1339676	B1339676 364724 MA
29	823	14.0	544	10	BF705690	BF705690 243606 MA
30	807	13.8	590	14	W48204	W48204 mc87h11.r1
31	782.5	13.4	867	9	AA755769	AA755769 vv35a11.r
32	781.5	13.3	554	10	BF191625	BF191625 239261 MA
33	762	13.0	767	12	BT734179	BT734179 603351465
34	761	13.0	763	9	AJ399251	AJ399251 AJ399251
35	759.5	13.0	564	12	BM088287	BM088287 501721 MA
36	725.5	12.4	540	9	AW505513	AW505513 UI-HF-BN0
37	723	12.3	925	13	BQ683553	BQ683553 AGENCOURT
38	721	12.3	1097	12	BM455276	BM455276 AGENCOURT
39	720.5	12.3	526	10	BF191620	BF191620 239255 MA
40	709.5	12.1	593	9	AA881653	AA881653 vx21c02.r
41	705	12.0	501	10	BE623895	BE623895 ut14d09.y
42	704	12.0	696	14	CB246966	CB246966 UI-M-F10-
43	702	12.0	436	9	AW141115	AW141115 EST291147
44	698	11.9	453	9	AA023709	AA023709 mh77e02.r
45	695.5	11.9	462	9	AW408832	AW408832 UI-HF-BM0

ALIGNMENTS

RESULT 1
AK038268
LOCUS
DEFINITION Mus musculus 16 days neonate thymus cDNA, RIKEN full-length enriched library, clone:Al30091E14 product:Janus kinase 3, full insert sequence.
ACCESSION AK038268
VERSION AK038268.1 GI:26332476
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
DB 82 TCAGAGGAGAGGAGCCTGCATGTGTCTCTCTCCCGGGGACCTGGGCTCTCCAGCGA 141
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
DB 142 TTGTCAATCTCTTTTGGGAGCTACTTGGCTGAGGATTTATGTGTGCGAGCTGCCAAGGCC 201
QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
DB 202 TGTGGCATCTCTCCCTGTTTATCATTCGCTTTTGTCTCTGCCCACTGAGACTTCTCTTGC 261
QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
DB 262 TGGTTTCCCCCAAGCCACATCTTCTGCATAGAGGAGTGGACACTCAAGTCTTGGTCTAC 321
QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
DB 322 AGGCTACGCTTTATTTCCCTGACTGGTTTGGCTGGAGACATGTCCCGCTTTGGGCTG 381
QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
DB 382 CGCAAGATTTGACCAGTCCCATCTTGACTTACAGTTTTTAGAATCTCTTTGTCTCAG 441
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
DB 442 CACCCAGTGCCTGGTGGGCGCTCCCGGTGGGCTTTAGCATGAGGAGCAGGGA 501
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaArg 180
DB 502 GAGTTCTCTGAGCCTGGCCGTGTGGACTTGGCTGGCCAGATGGCTCGTGAGCAGGCCAGCGC 561
QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
DB 562 CCAGGAGAGCTCTCTGAAGACGCTCAGTTACAAGCCTGTCTCGCGCCAGCCTGGCGCAT 621
QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAla 220
DB 622 GTGATCCAGGCGCAGAACTTCGTGACACGCGGCGCATCCGACGACCCGTGTGTGGCG 681
QY 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLe 240
DB 682 CTGCG-CCGTGTGTCTGCGCCAGCGCCAGCGCTACGCGCTCATGGCCAAATATATTTT 740
QY 240 uAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260
DB 741 GGACCTGGAGCGGCTACATCCAGCGGCGCACCCAGGAGACTTCCGCTGGGGCTCCCGGG 800
QY 260 yAlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrpSerSe 280
DB 801 CCCCCAGGAGGAGCGGGGCTTCTGGCTGTGGCGGGGGAACCGCATCTCTGGAGCTC 860
QY 280 rAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIl 299
DB 861 CGGGGACCCAGGAGCTTTCCAGACCTTCTGTGACTTTCGGAAATCGTGGATGTGAGCAT 920
QY 299 eAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMe 319
DB 921 CAAGCAGGCCCCCAGCTGGGTCCGCGAGGGGAACACCGGCTGGTCACCGTCAACAGGAT 980
QY 319 tAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAl 339
DB 981 GAGCGGCCACATCTCTGGAAGCGGAGTTTCCGGGGCTGCCCTGAGGCGCTGTCTTCTGTGGC 1040
QY 339 aLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluVa 359
DB 1041 CCTTGTGGATGGTACTTCCGCTGATCTGCGACTCCAGGCAATTTCTGCAAGGAGGT 1100
QY 359 lAlaProProArgLeuLeuGluAlaAspValCysHisGlyProIleThrLeuAs 379
DB 1101 GCGCGCGCGCGCTCTCGAGGAGGAGCGGAGCTGTGCCATGGACCCATCACGTTAGA 1160
QY 379 pPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgAr 399

DB 1161 CTTTCCCATCCACAAGCTGAAGCGCGCTGGCTCCCTCCAGGACACCTACATTTCCGCGC 1220
QY 399 gSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyPr 419
DB 1221 CAGCCCGGAGGACTATGACAGCTTTCTTTACCGCGCTGCGTCCAGACTCTCTTTGGCC 1280
QY 419 oAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLe 439
DB 1281 CGACTACAGGGCTGCCCTCATCCGCGAGACCCAGCGGGGCTTTCTCCCTGGTTGGCCT 1340
QY 439 uSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuAr 459
DB 1341 CAGCCAGCCCCACAGAGGCTCGGGAGCTGCTTGACGCTCTGCGAAATCTTGGGCTGCG 1400
QY 459 gValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGluLysSe 479
DB 1401 AGTAGCGGTGTGCCCTGAACCTAACATCTCTGCTGCGCTGCCAGACCCCAAGAAAGTC 1460
QY 479 rAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSe 499
DB 1461 CAATTTGATCGTGGTGGAGGGGCTGCACCCCGCGCTGCCCTGGCTGCTCCCCGTC 1520
QY 499 rCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHi 519
DB 1521 CTGCTGTGCGCTGACACAGCTGAGCTTCCACACAATTCACACGACAGCTCGAGTGGA 1580
QY 519 sGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValVa 539
DB 1581 CGAGAACCTGGGTACCGGTTCTTTTACCAAGATCTTCGTGGCGCGAGCGGAGGTCT 1640
QY 539 lAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAs 559
DB 1641 GGATGTTGAGACACATGACTCGAAAGTCTCTCTGAAGGTATGGACTTCAGACATCGGA 1700
QY 559 nCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLe 579
DB 1701 CTGCATGGAGTCTTTCTGGAAGCGCAAGCTTGATGAGCCAAAGTATCTATCCCGCACCT 1760
QY 579 uValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTy 599
DB 1761 GGTGTTACTGCACGGCGTCTGCATGGCTGGAGACAGCATCATGCTGCAGAAATTTGTGTA 1820
QY 599 rLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLy 619
DB 1821 TCTAGAGCAATGACATGACTCTGCGCAAGCGTGGCCACCTGGTGTACGCAGCTGGAA 1880
QY 619 sLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuPr 639
DB 1881 ACTGCAGGTGACCAAGCAGCTGGCATATGCCCTTAACCTTGGAGGACAAAGGCCCTTC 1940
QY 639 oHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGlyAspGlyAsnPr 659
DB 1941 TCACGGCAACGCTCTCAGCACGGAAGGTGCTCTCGTGTCTGTAGGGGGGTGATGGAAATCC 2000
QY 659 oProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetIe 679
DB 2001 ACCTTTTCATTAAGCTGAGTGATCTCGGTGTAGTCCCCTGCTGCTGAGGCTGGAAATGCT 2060
QY 679 uThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLe 699
DB 2061 CACCCAGCAATACCTCGGTGGCCCCCGAATGTCTCCAGGAGGCTCAGACACTCGGCTT 2120
QY 699 uGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAl 719
DB 2121 GGAGGCTACAAGTGGGGCTTTGGAGCCACACGCTGGGAGGTGTTCAGCGGGGACCCGCG 2180
QY 719 aHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnIe 739
DB 2181 CCACATCACTCTCGCTGGAGCGCCCAAAAGCTGAAGTTCTATGAGGACCCAGGACAGCT 2240
QY 739 uProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPr 759

Db 2241 GCCCGCTCTCAAATGGACAGAACTGGCGGAGCTTATCACAGTGCATGCGCTATGATCC 2300

QY 759 oGlyArgArgProSerPheArgAlaLeuLeuArgAspLeuAsnGlyLeuLeuThrSerAs 779

Db 2301 TGGCGGCGCCCTCTTCGAGCTATCCCTCAGAGACTCAACGGCTCTATTACATCAGA 2360

QY 779 pTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysVa 799

Db 2361 TTACGAGCTCTCTCAGACCCACACCTGGCATCCGAGTCTCGAGATGAGCTGTGC-- 2418

QY 799 lAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLy 819

Db 2419 -GGTGGCGCCAGCTCTATGCTGCCAGGACCCCGCATATTCAGGAGAGACACCTTAA 2477

QY 819 sTyrIleSerLeuLeuGlyLysGlyAnPheGlySerValGluLeuCysArgTyrAspPr 839

Db 2478 GTACATCTCTTGTGGCGAAGGCACTTGGCAGCGTGAGCTGCGCTATGACCC 2537

QY 839 oLeuGlyAspAnThrGlyProLeuValAlaValLysGlnLeuHisSerValProAs 859

Db 2538 CCTGGGGGCAATACGGGACCCCTGTGGCAGTGAACAGCTACAGCAGCGGGCCAGA 2597

QY 859 pGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheI 879

Db 2598 CCAGCAGAGGAGCTTCCAGCGGAGATTTCAGATCCTTAAGGCTCTGCACAGCGACTTCAT 2657

QY 879 eValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetG 899

Db 2658 CGTCAAGTACCGGGAGTCACTATGGCCAGGTCGCCAGAGCTCGGTTGGTGGTGA 2717

QY 899 uTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly---LeuHisTh 918

Db 2718 GTACCTGCCAGCGCTCGCTCGCAGACTTCTCGAGCGCATCGCGCGCTGCACAC 2777

QY 918 rAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaAr 938

Db 2778 CGACCGCTACTGCTGTTCGCTGGCAGATTCGACAGGCGATGAGTACCTGGTGGCGG 2837

QY 938 gArgCysValHisArgAspLeuAlaAlaArgAsnIleuValGluSerGluAlaHisVa 958

Db 2838 CCGCTGCTACACCGTACCGTGGCTGGCGGCAATCTTGTGGAGAGCGAGGCTCATGT 2897

QY 958 lLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVa 978

Db 2898 GAAGATCGCGACTTCGGCTCGTACGCTGCTGCCCTCGGAAGGACTACTACGTGT 2957

QY 978 lArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnI 998

Db 2958 CCGCGAGCTGGCCAAAGCCCATCTTTTGGTATGCCCGGAGTCCCTATCTGACACAT 3017

QY 998 ePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTy 1018

Db 3018 CTTCCTCCGCAATCTCAGCTGTGGAGTCTCGGAGTGGTGTGTACGAGCTCTTACCTA 3077

QY 1018 rCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgG 1038

Db 3078 CTGCGACAGAGCTGCAGCCCATCGCTAGTTCCTGAGCATGATGGGCTCGAGCGTGA 3137

QY 1038 uGlyProProLeuCysArgLeuLeuGluLeuAlaGluArgArgLeuProProPr 1058

Db 3138 AGGACCCCGCTCTGCGGCTCTCGGAGTGTGTCGAGAGGCGGAGCGCTCCACACCC 3197

QY 1058 oProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGluProHi 1078

Db 3198 TCCCACTTGGCCCAAGGTTACAGAGTCTATCAGCTGTGCTGGGCGGCCACCGCGCA 3257

QY 1078 sAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgPr 1098

Db 3258 CGACCGGCCAGCTTGGCACCTTGGAGCCCCCAGCTGGACGCGCTGTGGGTGGAGACC 3317

QY 1098 oGly 1099

Db 3318 CGGA 3321

RESULT 2

AK088365
LOCUS
DEFINITION

AK088365 3781 bp mRNA linear HTC 05-DEC-2002
Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN
full-length enriched library, clone:E430014A10 product:Janus kinase
3, full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK088365 GI:26353349
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

REFERENCE

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20493174
11042159

REFERENCE

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11078661

TITLE

JOURNAL
MEDLINE
PUBMED
AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hata, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L.M., Stauble, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, N., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Tayo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohseki, S.
and Hayashizaki, Y.

TITLE

JOURNAL
MEDLINE
PUBMED
AUTHORS

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3781)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sugabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
URL: http://location.qualifiers

FEATURES

source

```
1. .3781
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:E430014A10"
/db_xref="taxon:10090"
/clone="E430014A10"
/cell_type="thymic cells"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"
63. .3362
/notes="unnamed protein product: Janus kinase 3
(MGI:MGJ:99928, GB|NM_010589, evidence: BLASTN, 98%,
match=3715)
putative"
/codon_start=1
/protein_id="BAC40305.1"
/db_xref="GI:26353350"
/translation="MAPPSEETPLIPQRSCSLSSSEAGALHVLPPRGPPQRLSFS
FDYLAEDLCVRAAKACGLIPVHSLFALATEDSCWFPFPHICLEDVDVTVYRL
RYFPDWFLETCRHLRGLKRLDLSAILDLHLEHFAQHRSDLVSRFLVGLSMKEQG
EFLSLAVLQAWAREAQGPGLLTVSYKALPSPSRDVIQGNFVRRIRRTV
LALRRVACQADRYALMAKYILDLERLHPAATTETPRVGLPGAEPGLLRVADNGI
SNSSGDELFQTFDFPEIVDYSIKOAPRVGAGHRLVTVMDGHILAEFPGLE
ALSFLVADGYFRLICDSRHYCKEVAPEPRLEERELCHGITLDPALHKAAGSL
PGTYILRRSQDIFLLTACVQTPIDYKGLIRKQDPGSAFSLVLSQPSUREL
LAAWNSGLVDGAALNLSCCAPRPEKSNLIVRRGCTPAPAGSCSCCALQTSF
HTPTDLSLWHENLHGSGFTKPRGRREVVDGETHDSEVLKVMDSHRNOMESFLE
AASLSQVSPHPLHLHGVCMAGDSIMVOEFVVLGAIMDLRKRLHVSAMKLVTK
QLAYALNYLEDKGLPHNGVSARKVLLAREGGDGNPPFIKLSDPGVSTVLSLEMLTDR
IPWAPCECLOEAQTLGLEADKWGFGATTWEVSGGPAHTISLEPAKLFYEDQGLP
ALKWTELAGLITQCMAYDPPRRSPFRALIRDLNGLITSYELLSDDTPGIPSPRDEL
GAQLVACQDPAIFERHLKYISLLGKNGFSGVELCRDPLDNGTPTGVAVKLOHSG
PQQRDFQREIQTILKALHSDFTVKYRGVSYPGRQSLRLVMEYLPSCGLRDLQRHRA
RLHTRDLFLFAGWICKMVLYGARCVHRDLAARNLIVSEAHVIAIDFGLAKLPLG
KDYVVRPEQSPFWAPESLSDNIFSQSDWMSGVVLYELFYCDKSCSPSBEFL
SMWPEREGFPICLLELLAEGRRLPPFTCTPTEVQELMQLCWAFSPHDPAPFGLTSP
QLDALWRGRPG"
3764. .3769
/notes="putative"
```

CDS

```
polya_site 3781
BASE COUNT 707 a 1193 c 1106 g 775 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3781
Score: 5650.00 Matches: 1077
Percent Similarity: 97.91% Conservative: 2
Best Local Similarity: 97.73% Mismatches: 18
Query Match: 96.42% Indels: 6
DB: 11 Gaps: 4
US-09-397-967A-16 (1-1099) x AK088365 (1-3781)
QY 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
Db 63 ATGGCACCTCCAAAGTGGAGGAGACCTCTGATCCCTCAGCGCTCTTGCAGCCTCTCATCC 122
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
Db 123 TCAGAGCGAGGAGCCCTGCTGATGCTCTCTCTCCCGGGGAGCTGGCGCTCCCGACGA 182
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
Db 183 TTGTCATTCTCTTTTGGGACTACTTGGCTGAGGATTATGCTGCGAGCTGCCAAGGCC 242
QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db 243 TGTGGCATCTCGCTGCTTTTATCATCTGCTTTTCTGCTTGGCCACTGAGGACTTCTCTTGC 302
QY 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db 303 TGGTTTCCCGCAAGCCACATCTTCTGCATAGAGGAGTGGACACTCAAGTCTTGGTCTAC 362
QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
Db 363 AGGCTACGCTTTTATTTCCCTGACTGTTGGCTGGAGACATGTACCGCTTTGGGCTG 422
QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
Db 423 CGCAAGATTTGACCAGTGCATCTTGTACTTACAGCTTTTAGAACATCTCTTTGCTCAG 482
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
Db 483 CACCGCAGTGACTGTGTGAGTGGCGCCCTCCCGGTGGGCTTACGATGAAGGAGGAGGGA 542
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
Db 543 GAGTTCTCTGAGCTGGCGCTGTGGACTTGGCCAGATGGCTCTGTGAGGAGGCCAGCGC 602
QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
Db 603 CCAGGAGAGCTGCTGAAGCGGTGAGTACAAAGCCGTCTGCGGCCAGCGCTGCGCGAT 662
QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAla 220
Db 663 GTGATCCAGGGCCAGAACTCTGTGACACGACGCGCATCCGAGGAGCCGTGTGTTGGCG 722
QY 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLe 240
Db 723 CTGGC-CCGTGTGTGCTGCTGCGACGCCACCGCTACGCGCTCATGGCCCAAGTATATTT 781
QY 240 uAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260
Db 782 GGACCTGGAGCGCTACATCCAGCGGCCACACCGAGACCTTCCGTGTGGGCTCCCGGG 841
QY 260 yAlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrSerSe 280
Db 842 CCGCCAGGAGGAGCGGGGCTTCTGCTGTGTGGCGGGGACCAACGGCATCTCTCGAGGCTC 901
QY 280 rAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIl 299
```

polyA_signal

Db 902 CGGGACCGAGAGCTTTTCCAGACCTTCTGTGACCTTTCCGGAAATCGTGGATGTGACGAT 961
Qy 299 eAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMe 319
Db 962 CAAGCAGGCCCCACGTGTGGGTCCCGCAGGAGGAAACACCGGTGTGACCCGTACACAGAT 1021
Qy 319 tAspGlyHisIleLeuGlnAlaGluPheProGlyLeuProGluAlaLeuSerPheValAl 339
Db 1022 GGACGGCCACATCTCTGGAGCGGAGTTTCCGGGCTGCTCCAGGGGTGTCTCTGTGGC 1081
Qy 339 aLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluVa 359
Db 1082 CCTTGTGATGGGTACTTCCGCTGATCTCGACTCCAGGCATTATTCTGCAAGGAGT 1141
Qy 359 lAlaProProArgLeuLeuGlnGluAlaAlaAspValCysHisGlyProIleThrLeuAs 379
Db 1142 GGCCTCCGCGGTGTCTGGAGGAGGAGCGGAGCTGTGCCATCGACCCATCAGCTTAGA 1201
Qy 379 pPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgAr 399
Db 1202 CTTTGCATCCACAAGCTGAAGCGCGTGGCTCCCTCCAGGACCTACATTTCTCCGCG 1261
Qy 399 gSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyPr 419
Db 1262 CAGCCCGCAGGACTATGACAGCTTCTTCTTACCGCTGCTGCTCAGACTCCTCTTGGGCC 1321
Qy 419 oAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLe 439
Db 1322 CGACTACAAGGGCTGCCTCATCTCCGCCAGGACCCAGCGGGCTTCTCCCTGGTGGCT 1381
Qy 439 uSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuAr 459
Db 1382 CAGCCAGCCCCACAGAAGCTTCGGGAGCTGCTGTGACCC---TGGAAATTCTGGGCTCG 1438
Qy 459 gValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaProArgProLysGluLysSe 479
Db 1439 AGTAGCGGTGCTGCCCTGAACCTTAACATCTGCTGCTGCCCTCCAGACCCCAAGGAAAGTC 1498
Qy 479 rAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSe 499
Db 1499 CAATTGTATCTGTGTGGAAGGGCTGCACCCCGCGCTGCCCTGCTGCTCCCGCTC 1558
Qy 499 rCysCyAlaAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHi 519
Db 1559 CTGCTGTGGCTGACACAGCTGAGCTTCCACACATTTCAACGACGACCTGGAGTGCA 1618
Qy 519 sGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValVa 539
Db 1619 CGAAACCTGGGTCACTGCTTTTACCAAGATCTTCCGTGGCCGCGAGGAGGTGCT 1678
Qy 539 lAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAs 559
Db 1679 GGATGGTGAGACATGATGACTCGGAAGTCTCTCTGAAGGTCTGGAATCCAGACATCGGAA 1738
Qy 559 nCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisIle 579
Db 1739 CTGATGGAGTCTTCTTGGAGCGCGAGCTTGTATGAGCCAGTATCTTACCCGCACT 1798
Qy 579 uValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTy 599
Db 1799 GGTGTACTGCACGCGCTGCTGCATGGCTGGAGACAGCATCATGTTGTCAGGAATTTGTGTA 1858
Qy 599 rLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLy 619
Db 1859 TCTAGGACCAATTGACATGTACTCTCGCAAGCGTGGCCACCTGTGTGTCAGCCAGCTGGAA 1918
Qy 619 sLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuPr 639
Db 1919 ACTGCAGGTGACCAAGCAGCTGGCATATGCCCTTAATCTTGTGGAGGACAAAGGCTTCC 1978
Qy 639 oHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyAspGlyAsnPr 659
Db 1979 TCACGGCAACGTCTCAGCAGGAAGGTGCTCCTGCTGCTGAGGGGGGTGATGGGAATCC 2038

Qy 659 oProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLe 679
Db 2039 ACCTTTTCATTAAAGCTGAGTGATCTCTGTGTGTCAGTCCACCTGTGTGAGCCCTGGAAATGCT 2098
Qy 679 uThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLe 699
Db 2099 CACCGACAGAATACCTCTGGGTGGCCCGGAATGTCTCAGGAGGCTCAGACACTCGGCTT 2158
Qy 699 uGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAl 719
Db 2159 GGAGCTGTGACAAGTGGGCTTTTGAGGCCACACCTGGAGGTGTTCAGCGGGGACCCGC 2218
Qy 719 aHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLe 739
Db 2219 CCACATCACCTCTGTCGAGCCCGCCCAAAAGCTGAGTTCTATGAGGACCCAGGACAGCT 2278
Qy 739 uProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPr 759
Db 2279 GCCCGCTCTCAATGGACAGAACTGGCGGAGCTTATCACAGTGCATGGCTATGATCC 2338
Qy 759 oGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs 779
Db 2339 TGGCGCGCGCCCTCTTCCAGCTATCTCAGAGACTCAACGCGCTCATTTACATCAGA 2398
Qy 779 pTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysVa 799
Db 2399 TTACGAGCTCTCTCAGACCCACACCTGGATCCCGAGTCTCTCGAGATGAGCTGTC-- 2456
Qy 799 lAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLy 819
Db 2457 -GGTGGCGCCAGCTCTATGCTGCCAGGACCCCGCCATATTCAGGAGAGACACCTTAA 2515
Qy 819 sTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspPr 839
Db 2516 GTACATCTTCTTGTGGCAAGGCAACTTTGGCAGCGTGGAGTGTGCCCTATGACCC 2575
Qy 839 oLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAs 859
Db 2576 CCTGGGAGACATACGGGACCCCTGTGTGGCAGTGAACAGCTACAGCACAGCGGCCAGA 2635
Qy 859 pGlnGlnArgAspPheGlnArgGluIleGlnIleLysAlaLeuHisSerAspPheI 879
Db 2636 CCAGCAGAGGACTTCCAGCGGAGATTTCAGATCCTTAAGCTCTGCACAGCGACTTCAT 2695
Qy 879 eValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGl 899
Db 2696 CGTCAAGTACCGGGAGTCACTATGGGCCAGGTGCGCAGAGCTTCCGCTTGGTATGGA 2755
Qy 899 uTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGly---LeuHisTh 918
Db 2756 GTACTGCCAGCGCTGCTTCCGAGACTTCTGCAGCGCCATCGCGCGCTGCACAC 2815
Qy 918 rAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaAr 938
Db 2816 CGACCGCTACTGCTGTTCGCTTGGCAGATCTGCAAGGCGATGGAGTACCTGGGTGCGCG 2875
Qy 938 gArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisVa 958
Db 2876 CCGTGGCTACACCGTACCTGCTGCGCGCAACATCTTGTGGAGAGCGAGGCTCATGT 2935
Qy 958 lLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVa 978
Db 2936 GAAGATCGCGACTTCGGCTCGTAACTGCTGCCCTCGTGGGAAGGACTACTACGTGT 2995
Qy 978 lArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnIl 998
Db 2996 CCGCAGGCTGGCCAAAGCCCATCTTTTGTGTATGCCCGGAGTCCCTATCTATCGAACAT 3055
Qy 998 ePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTy 1018
Db 3056 CTTCTCCCGCAATCTGACGTGTGGAGCTTCGGAGTGTGTGTACGAGGCTCTTCCACCTA 3115

QY 1018 rCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgG1 1038
 Db 3116 CTGCGACAGAGCTGAGTCCATCCGCTGAGTTCTTGAGCATGATGGGCTGAGCGTGA 3175

QY 1038 uGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgArgLeuProProPr 1058
 Db 3176 AGGACCCCGCTGCTGCGCTCTGAGCTGCTGGCAGAGGGCGAGCGCTCCACCACC 3235

QY 1058 oProThrCysProThrGluValGlnGluLeuMetGlnLeuLeuLeuAlaProGluProHi 1078
 Db 3236 TCCCACTGCCCCACAGAGTTGAGAGCTCATGAGCTGTGCTGGCGCCGACCCGCA 3295

QY 1078 sAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgPr 1098
 Db 3296 CGACCGGCGAGCTTCCGACCTTGGCACCCTGAGCCCCAGCTGGAGCGCTGTGGCTGGAAGACC 3355

QY 1098 oGly 1099
 Db 3356 CGGA 3359

RESULT 3
 BC053021
 LOCUS BC053021 4462 bp mRNA linear HTC 05-JUN-2003
 DEFINITION Mus musculus cDNA clone IMAGE:5709138, containing frame-shift errors.

ACCESSION BC053021
 VERSION BC053021.1 GI:31418517
 KEYWORDS HTC.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 4462)
 AUTHORS Strausberg,R.D., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Frange,C., Roha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,J.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 12477932
 22388257

REFERENCE 2 (bases 1 to 4462)
 AUTHORS Strausberg,R.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapsb@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 114 Row: h Column: 3
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9055371
 This clone has the following problem: frame shifted.

FEATURES

source
 1. 4462
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5709138"
 /tissue_type="Brain, mouse 15.5 dpc"
 /clone_lib="NIH BMAP_EXO"
 /lab_host="DH10B"
 /note="Vector: pYX-ASC"

BASE COUNT 986 a 1261 c 1268 g 947 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,27e-116 Length: 4462
 Score: 1447.00 Matches: 411
 Percent Similarity: 45.56% Conservative: 138
 Best Local Similarity: 34.11% Mismatches: 396
 Query Match: 24.69% Indels: 262
 DB: 11 Gaps: 35

US-09-397-967A-16 (1-1099) x BC053021 (1-4462)

QY 2 AlaProProSerGluThrProLeuLeuProGlnArgSerCysSerLeuSerSerSer 21
 Db 228 GCAGATGGCACAGAGCTCAGCCCTCGTGTGCC----- 260

QY 22 GluAlaGlyAlaLeuHisValLeuLeuProProArgGlyPro---GlyProProGlnArg 40
 Db 261 ---ACAGGATGCTGTGATGTGTACTGTGCAATGGCCCTGGGCGGGAGCCCTGG 317

QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaLysAla 60
 Db 318 GTCACTTCAGCCAGACATCTGTGCTGCAGAGGAGTCTGCATCCACATCGCACAAA 377

QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
 Db 378 GTCCGATCACTCCACCTGCTTGAATCTCTCGCACTCTACAATGCACAGGCTAAGTTC 437

QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
 Db 438 TGGCTGCCCCCAACCATATT-----CTGGATACATCCCAAGACATGAACCTTATTTT 491

QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGlu-----ThrCysHis 116
 Db 492 CGAATCAGGTTTACTTCGGAACTGGCATGCAATCCCGAGGAGCCAGCTGTATAC 551

QY 117 ArgPheGlyLeuArgLysAspLeuThrSerAla-----Ile 128
 Db 552 CGGTGTGTTCCTCCAGGGGCGAGACTTCTCAGACCGGGCAGAGCAAGGTGTACAGCTC 611

QY 129 LeuAspLeuHisValLeuGluHisLeuPheAlaGlnHisArg-----SerAsp 144
 Db 612 TTGGACTTCGCTCGTTGAATACCTTTTGTAGCGGGAAGCATGATTCATGAACGAT 671

QY 145 LeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGln----- 159
 Db 672 GTGGTGTCTCTCGG-----GACCTGTCTAGGAGGAGATCCACCACCTTTAAG 722

QY 160 GlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGln 179
 Db 723 AATGAGAGCTTAGCATGCGCTTCTGCAACCTGTGTACCTTGTCTCAGCGGCGCTC 782

QY 180 ArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArg 199
Db 783 CCCTGGAGGAGTGGCCAGAGATCAGCTTCAAGAACTGCATCCCTCATCTCCGA 842
QY 200 AspValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeu 219
Db 843 CAGCACATCCGGCAGCACAATGTCTCACAGCCTGGGGTCCACAGA--GTCTCCGC 899
QY 220 AlaLeuLeuProCysGlyArgLeuProGlyArg-----ProTyrAlaLeuMetAlaLys 237
Db 900 CGCTTCCTGGGGCTCCCG--CCYGGCCACTCTCCAGCAGGTTGTGTGATGGAAG 956
QY 238 TyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGly 257
Db 957 TACTTGGCTACCTGGAGCGCTGGCTCCGGC----- 989
QY 258 LeuProGlyAlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePro 277
Db 990 -----TTTGGCTCAGAGCGCATACCT 1010
QY 278 TrpSerSerAsnAspGluLeuPheGlnThrPheCysAspPheProGluIleValAspVal 297
Db 1011 -----GTGTGTCATCTG 1022
QY 298 SerIleAsn-GlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValTh 317
Db 1023 GAGGTCTGGCAGACCCGAGAGGACCCCTGCTATACCCAGA-----ACAGTGGG 1073
QY 317 rArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPh 337
Db 1074 CAGACCGCTGGGGATCCAGGCCAGAGCTGCTTCCGGGCTCCACCCAGAGTACTG 1133
QY 337 eValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLy 357
Db 1134 GTGACAGGCACCGAGGATCCAGTGGCATCCGCTGCAGACCCAGGGATCCGTTTGTACA 1193
QY 357 sGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIleTh 377
Db 1194 G-----GCCAAGCTGTGCCAGAGGAC----- 1215
QY 377 rLeuAspPheAlaIleHisLysLeuLysAlaGlySerLeuProGlyThrThrIleLe 397
Db 1216 -----GGCCTCTACCTGAT 1229
QY 397 uArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLe 417
Db 1230 TCAGTGGAGCACCGCCACCTGCACCGCTGATCTTACCGTGGCCCATCGAAACCCCGC 1289
QY 417 u-----GlyProAspTyrLysGlyCysLeuIleArgGlnAspPro----- 430
Db 1290 TTTTCAGTAATGGCCCT-----CGGGGCTGGCCTGGAAAGTTCCCCCATCACACAGA 1343
QY 431 -SerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLe 450
Db 1344 GCCTGGAGCCTTTGTCTGGATGGCTGGGGCGCTCTTTGCCACCTTGGGGGACCTTCG 1403
QY 450 uAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCy 470
Db 1404 GCTCGCCTTTCAGGGCTGCTGCTTCGGGCCCGGTGATGACTGTTTCCCTTCACCACTG 1463
QY 470 sCysAlaProArgProLys----- 476
Db 1464 CTGCTGCCCCCGCCAGAGGTTAGTAGAGAGTGGGAAACACAGGGTAGGATGAGTCT 1523
QY 477 -----GluLysSerAsnLe 481
Db 1524 GTACCTGAAGCCAGCCACAGGGCGATGTCCACTCCTGTGCTTAGAAATCTCCAACT 1583
QY 481 uIleValValArgArgGlyCysAsnProAlaProAla----- 493
Db 1584 GTCTATATGCGGGGTCTAG--GGCCACACCCCGCTCTCAACCTCAGTCAGTCAGCT 1642
QY 494 -----ProGly----- 495

Db 1643 TCCACAGGGTTCCACAGGATGAATACCCAGGAGTGGTGTCTCTGGCAGGACGTGCAGA 1702
QY 496 -----CysSerProSerCysCysAlaLeuThrGl 505
Db 1703 CTCTGGGGGGGGGGGTGCACCTGTATCCAGTGCACCCACACTGC----- 1748
QY 505 nLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGlyHisGl 525
Db 1749 -ATTCTGTACCACCTTCTTACAGCTG-----TCCCACTTGGGCCAAGG 1789
QY 525 ySerPheThrIlePheArgGlySerArgArg-----GluVa 538
Db 1790 CACAAGACCAATGTATGAGCGCTTCTAAGAGTGGGAGGCCGCCGATGAGGGCAAGT 1849
QY 538 lValAspGly-----GluThrHisAspSerGluValLeuLe 550
Db 1850 GGACAATGGATGTCCCTCTGAGCTGTGGGACTAGTGGGAGCAGCTTCGAGTGGTCT 1909
QY 550 uLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLe 570
Db 1910 CAAGTTCTCGACCCCATGCAC----- 1931
QY 570 uMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAs 590
Db 1932 -----CATGATATCGC-CTT----- 1945
QY 590 pSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysAr 610
Db 1946 -GATATCATTTGTGACAGAAATTCGAGAACATGGTCCCTTGGATGTGTGTTACGGCGACA 2004
QY 610 gGlyHisLeuValSerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLe 630
Db 2005 GAGGGGCCAAGTGGCCCATGACCTGGAGATGGTGTGCTCAGCAGCTGGCCAGCGCCT 2064
QY 630 uAsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLe 650
Db 2065 CAGTACTGTGAGGACAAAGATCTGGTTCACGGAAATGTATGTGGCCGGAACATCTCTCT 2124
QY 650 uAlaArgGluGly--GlyAspGlyAsnProProPheIleLysLeuSerAspProGlyVa 669
Db 2125 GGCTCGCCTGGGGTGGAGGAGGTACCAACCCCTTCATCAAGTAACTGATCTCTGGTGT 2184
QY 669 lSerProThrValLeuSerLeuGluMetLeuThrAspArgIleProThrValAlaProGl 689
Db 2185 GGGCCAGGGTCCCTCTCTCCAGGGAAGAGCGGGTGGAGCGCATCCCTCGACAGCTCCCGA 2244
QY 689 uCysLeuGln---GluAlaGlnThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAl 708
Db 2245 GTGCTGTCTGGAGGGACCATAGCTTGGGTACTCGCCACGGACATGTGGGGCTTTGGTGC 2304
QY 708 aThrThrTrpGluValPheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaLy 728
Db 2305 CACCTTCTTGAGATCTGCTTTGATGGGAGCGACCCCTCGAGGGTCTGTCTCTCTGA 2364
QY 728 sLysLeuLysPheTyrGluAspGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAl 748
Db 2365 GAAAGACGGTTCTACACAAAGAAACATCAGCTGCCTGAACCTCATCCCCAGAGCTGGC 2424
QY 748 aGlyLeuIleThrGlnCysMetAlaTyrAspProGlyArgArgProSerPheArgAlaI 768
Db 2425 CACACTCAACCGCCAGTGCCTATGAACCAACAGCAGCGGCCATCTTCCGACCAT 2484
QY 768 eLeuArgAspLeuAsnGlyLeuLeuThrSerAspTyrGluLeuSerAspProThrPr 788
Db 2485 TTTGGGAGCTTCACAGGCTG-----CAGCCAGAAATCTAGTCGCGACTTCGGCTGT 2538
QY 788 oGlyIleProSerProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGl 808
Db 2539 GAACCTCAGACTCACCA-----GCATC 2559
QY 808 nAspProAlaIlePheGluGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAs 828
Db 828 -----GCATC 2559

```
Db 2560 AGACCCCACTGTTTTCACAGCGCTATTTGAAAAGATCCGGGATTTGGGTGAGGCTCA 2619
QY nPheGlySerValGluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuVa 848
Db 2620 CTTTGGCAAGTACGCTGCTAGCTAGCCCAACCAATGACCGCACTGCGGAGATGGT 2679
QY lAlaValLysGlnLeuGlnHisSerValProAspGlnGlnArgAsp---PheGlnArgG 867
Db 2680 GGCCTGTGAAGCCCTGAAGGAAGGTCGGTCCCGAGCTCGCTCAGCTGGCAGCGGA 2739
QY nIleGlnIleuLysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTy 887
Db 2740 GATCGAGATCCTCGGAGCGTGTACCAACGAAATATTGTCAAGTATAAAGGCTCTGTGA 2799
QY rGlyProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuAr 907
Db 2800 GGACCAAGGAGAGAAGTCTGTACAGCTGGTCATGGAGTATGCTCCTCTGGCGACGCTCCG 2859
QY gAspLeuLeuGlnArgHis---ArgGlyLeuHisThrAspArgLeuLeuLeuPheAlaTr 926
Db 2860 AGACTACCTGCCAAGGCACTGCTAGGCTG-----GCGCAGCTCCTGCTGTTGCCCA 2913
QY pGlnIleCysLysGlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAl 946
Db 2914 GCAGATCTCGGAGGCGATGCCCTACCTGCACGCTCAGCACTACATTCACCGAGACCTCGC 2973
QY aAlaArgAsnIleuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAl 966
Db 2974 CGCGCGCAGCTGCTGTGACACGACGAGCTGGTCAAGATTGGAGACTTTGGCCTAGC 3033
QY aLysLeuLeuProLeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerProIl 986
Db 3034 CAAGGCTGTACCTGAAGGCCACGAGTACTACCGAGTGGCGAGGACGGGACAGCCAGT 3093
QY ePheTrpTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTr 1006
Db 3094 GTTCTGTATGCCCCAGAAATGCTCGAAGGAGTGCATAATTTTACTATGCTATGCTGCTG 3153
QY pSerPheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSe 1026
Db 3154 GTCTTTCGGGGTAACCTCTGTATGAGTTGTGATACATCTGACTCTAACCCAGAGTCCCCA 3213
QY rAlaGluPheLeuArgMetMetGlyProGluArgGluGlyProProLeuCysArgLeuLe 1046
Db 3214 TATGAATTCACCGAGCTCATCGCCATACCCAGGCGCAGATGACCGTGTGAGGCTCAC 3273
QY uGluLeuLeuAlaGluGlyArgArgLeuProProProProProProProProProProPro 1066
Db 3274 AGAGCTGCTGGAACGAGGAGAGAGGCTGCTCGGCTCGCAGGTGTCCCTGTGAGATCTA 3333
QY nGluLeuMetGlnLeuCysTrpAlaProGluProHisAspArgProAlaPheAlaThrLe 1086
Db 3334 TCACCTCATGAAGAATCTGGGAGACAGAGGCTGCTCCGCGCCACCTTCCAGAAATCT 3393
QY uSerProGlnLeu 1090
Db 3394 TGTGCCCATCTCTC 3406

RESULT 4
BI411962
LOCUS 602966375F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5121786 5',
DEFINITION mRNA sequence.
ACCESSION BI411962
VERSION BI411962.1 GI:15172885
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Eutelestomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 971)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
```

JOURNAL
COMMENT

Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1297 row: e column: 19
High quality sequence start: 25
High quality sequence stop: 785.

FEATURES

source

1. .971
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5121786"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NCI_CGAP_Lu33"
/notes="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGGCTCTGTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 216 a 274 c 279 g 202 t
ORIGIN

Alignment Scores:

Pred. No.: 1,94e-111 Length: 971
Score: 1380.00 Matches: 287
Percent Similarity: 89.30% Conservative: 5
Best Local Similarity: 87.77% Mismatches: 27
Query Match: 23.55% Indels: 10
DB: 12 Gaps: 2

US-09-397-967A-16 (1-1099) x BI411962 (1-971)

QY 475 ProLysGluLysSerAsnLeuValValArgArgGlyCysAsnProAlaPro 494
Db 6 CCAAGGTCAGAGTCATTTTGTGCTGGTGGAGGGGCTGCACCCCGCGCTGCCCT 65
QY 495 GlyCysSerProSerCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAsp 514
Db 66 GGCTGTCCCGCTCTGCTGTGCTGCACACAGCTGAGCTTCACACAATTCACCGGAC 125
QY 515 SerLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySer 534
Db 126 AGCTGTGAGTGGCAGAGAACCTGGGTACGGTTCCTTTTACCAAGATCTTCCGTGCGCG 185
QY 535 ArgArgGluValValAspGlyGluThrHisAspSerGluValLeuLysValMetAsp 554
Db 186 AGCGGGAGGTCGTGGATGTGAGACACATGACTCGAAGTCTCTCTGAAGGTTCATGGAC 245
QY 555 SerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnVal 574
Db 246 TCCAGACATCGGAATCTGCATGGAGCTTTTCTGGAAGCGCGCAAGCTTGATGAGCAAGTA 305
QY 575 SerTyrProHisLeuValLeuHisGlyValCysMetAlaGlyAspSerIleMetVal 594
Db 306 TCTACCCGCACTGCTGTACTGACGCGCTGTGATGCGTGGAGACAGCATCATCGGTG 365
QY 595 GlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuVal 614

366 CAGGAATTTGTATCTAGGAGCAATTTGACATGTACCTGGCGAAGCGTGGCCACCTGGT 425
 615 SerAlaSerTrpIysLeuGlnValThrIysGlnLeuAlaValAlaLeuAsnTyrLeuGlu 634
 426 TCAGCCAGCTGGAAACTGCAGGTGACCAAGCAGTGGCATATGCCCTTAACCTATTGGAG 485
 635 AspLysGlyLeuProHisGlyAsnValSerAlaArgIysValLeuLeuAlaArgGluGly 654
 486 GACAAAGGCTTCTTCACGGCAAGCTCTCAGCAGGAAGTGTCTCTGGCTCGTAGGGG 545
 655 GlyAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeu 674
 546 GGTGATGGGAATCCACCTTTTCATTAAAGCTGAGTGATCCTGGTGCAGTCCCACTGTGCTG 605
 675 SerLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAla 694
 606 AGCTGGAAATGCTCACCGCAGAAATACCTGGGTGGTCCCGCAATGTCTCCAGAGGCT 665
 695 GlnThr-LeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPh 714
 666 CAGACAACTGCTGTGGAGGTGACAGTGGGGCTTTGGAGCCACCAGTGGGAAGTGT 725
 714 eGlnArgGlyProAlaHisIleThrSerLeuGlu-ProAlaLysLysLeuLysPheTyrG 734
 726 CAGCGGGGACCCGCCCAATCACCTCGCTGGAGGCCCGCCCAAAAGCTGAAGTCTCTATG 785
 734 LuAspGlnGlyGlnLeuProAlaLeu---LysTrpThrGluLeuAlaGly-LeuIleTh 752
 786 AGGACCAAGGACCAAGTCCCGGTCTCAATGGCAGACAGAACTGGCCGGGAACCTTATCAC 845
 752 rGlnCys-MetAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeu-ArgAsp 771
 846 ACAGTGCAATGGGTATGAATCCTGGCGGGCCCTTCATTCCGAGATATCCTTCAGAGAC 905
 772 LeuAsnGlyLeuIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIlePro 791
 906 CTCAGGCGCT-ATTACATTGGTTTACGAGCTGTCTTA-AAACCCCAAACTGGAT---CCC 960
 792 SerProArg 794
 961 AGTCCGAGA 969

RESULT 5
 BC027234
 LOCUS
 DEFINITION Mus musculus, Similar to Janus kinase 3, clone IMAGE:3489805, mRNA.
 ACCESSION BC027234
 VERSION BC027234.1 GI:20071020
 KEYWORDS HTC.

SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1625)
 Strausberg, R.
 Direct Submission
 Submitted (04-APR-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louleseg, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,

REMARK
 COMMENT

A.N., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
 Series: IRAC Plate: 35 Row: g Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarity but not
 identity to protein
 This clone has the following problem: retained intron.

FEATURES

source
 Location/Qualifiers
 1..1625
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:3489805"
 /tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
 old, gross tissue."
 /clone_lib="NCI CGAP Mam5"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 BASE COUNT 328 a 470 c 483 g 344 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,71e-109 Length: 1625
 Score: 1358.00 Matches: 281
 Percent Similarity: 75.59% Conservative: 7
 Best Local Similarity: 73.75% Mismatches: 11
 Query Match: 23.17% Indels: 84
 DB: 11 Gaps: 3

US-09-397-967A-16 (1-1099) x BC027234 (1-1625)

QY 1 MetalProProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
 DB 42 ATGGCACCTCCAAAGTGAGGAGACACCTCTGATCCCTCAGCGCTCTTGACGCTCTCATCC 101
 QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProArgGlyProGlyProGlnArg 40
 DB 102 TCAGAGGAGGAGGCCCTGCATGTGCTCTCTCCCGGGGACCTGGGCGCTCCCGAGCGA 161
 QY 41 LeuSerPheSerPheGlyAspTyrIleuAlaGluAspLeuCysValArgAlaAlaLysAla 60
 DB 162 TTGTCAATCTCTTTTGGGAGCTACTTGGCTGAGGATTTATGTGTGCGAGCTGCCAAGGCC 221
 QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
 DB 222 TGTGGCATCCTGGCCCTGTTTATCATTCGCTTTTGGCTTGGCCACTGAGGACTTCTCTTGC 281
 QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
 DB 282 TGGTTTCCCCCAGCCACATCTTCTGCATAGAGGAGCTGGACACTCAAGTCTTGTCTTAC 341
 QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
 DB 342 AGGCTACGCTTTTATTTCCCTGACTGGTGTGGCTGGAGACATGTACCGCTTTTGGGTG 401
 QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
 DB 402 CGCAAGATTTGACCAGTGCCATCTTGCATTACATGTTTAGAACACAGTCTTGTCTCAG 461
 QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
 DB 462 CACCGCAGTGACCTGGTGGTGGGCGCTCCCGGTGGGCTTTAGCATGAAGGAGCAGGGA 521
 QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
 DB 522 GAGTTCTCTGAGCCTGGCGCTGTGACTTGGCCAGATGTGCTCGTGAGCAGGCCCGCCG 581
 QY 181 ProGlyGluLeuLeuLysThrVal----- 188
 DB 582 CCAGGAGAGACTGCTGAAGACGGT-CAGGTGAGGGGCTCGGGCCACCCCTTCTGTGCTTTC 640


```

188 ----- 188
641 TTGGGCTCGGATCTAGGGCCCTATCGTTTCCCTTACTTGAACACATACCTCTTCCCA 700
188 ----- 188
701 ACTTGTGACAGCCACCTCTCCGTGAGCCTGCCCTGCGGAAAGGTCCCTTATGTGTCT 760
189 -----Ser 189
761 CGTGTCTCTTGGGGCCCTCCACCCCTAAGAGTTCGGCGTGCCTCCCTCCCTCCAGT 820
190 TTTTCTCTCTTGGGGCCCTCCACCCCTAAGAGTTCGGCGTGCCTCCCTCCCTCCAGT 820
190 TTTTCTCTCTTGGGGCCCTCCACCCCTAAGAGTTCGGCGTGCCTCCCTCCCTCCAGT 820
821 TACAAAGCTGTCTGCGGCCACCTGCGGATGTATCCAGGCGCAGAACTTCTGTGACA 880
210 ArgArgArgIleArgThrValValLeuAlaLeuLeuProCysGlyArgLeuProGly 229
881 CGCAGGCGCATCCGAGGACCGTGTCTTGGCGTGGC-CGGTGTGTCTGCGCTGCCAGGC 939
230 ArgPro-TyrAlaLeuMetAlaLysTyrIleLeuAspLeuGluArgLeuHisProAlaAl 249
940 CGACCGCTACGCGCTCATGCCAAGATATCTGGACCTGGAGCGGCTACATCCAGCGGC 999
249 aThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGluGluProGlyLeuLeuAr 269
1000 CACCACCGAGACCTTCCGTGTGGGCTCCGGGCGGCCAGGAGCGCGGGCTTCTGCG 1059
269 gValAlaGlyAspAsnGlyLeuProTrpSerSerAsnAspGluLeuPheGlnThrPheCy 289
1060 TGTGGCGGGGACACGCGATCTCTCTGGAGTCCGGGACCAG----- 1102
289 sAspPheProGluLeuValAspValSerIle---AsnGlnAlaProArgValGlyProAl 308
1103 -----GAGTGTCTGGGCTGGGCTTGAGAGCGGGGCTCCCGGGAGCGGTGGC 1152
308 a 308
1153 G 1153

RESULT 6
BX415899
LOCUS
DEFINITION
5-PRIME mRNA sequence.
ACCESSION
BX415899
VERSION
BX415899.1 GI:30641720
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1200)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9047.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0CAP008B09QPl&cluster=9047.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0CAP008B09QPl.
Location/Qualifiers
1..1200
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES
source

```

```

/ab_xref="taxon:9606"
/clone="CS0CAP008YM18"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized.
BASE COUNT 247 a 337 c 340 g 234 t 42 others
ORIGIN
Alignment Scores:
Pred. No.: 1,15e-96 Length: 1200
Score: 1214.50 Matches: 245
Percent Similarity: 83.59% Conservative: 25
Best Local Similarity: 75.85% Mismatches: 51
Query Match: 20.73% Indels: 6
DB: 13 Gaps: 1
US-09-397-967A-16 (1-1099) x BX415899 (1-1200)
QY 289 CysAspPheProGluLeuValAspValSerIleAsnGlnAlaProArgValGlyProAla 308
Db 89 TGCAGCTTCCAGAAATCGTAGACATTAGCATCAAGCAGCCCGCGCGTGTGGCCGGCC 148
QY 309 GlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPhe 328
Db 149 GGAGACACCGCTGCTGCTGCTTACAGGACAGACCAACAGATTTTAGAGCCGAGTTTC 208
QY 329 ProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuLeu 348
Db 209 CCAGGCTCCCGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 268
QY 349 CysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGluGlu 368
Db 269 ACGGACTCCAGCACTTCTTCTGCAAGAGGTGGCACCCCGCGAGGTGTGTGGAGNAGTG 328
QY 369 AlaAspValCysHisGlyProIleThrLeuAsp-PheAlaIleHisLysLeuLysAlaAl 388
Db 329 GCCGAGCAGTGCACGGCCCATCTCTGGACTTKGSSRVKCAACCAAGCTCAAGACTGG 388
QY 388 aGlySerLeuProGlyThrTyrIleLeuArgSerProGlnAspTyrAspSerPheLeu 408
Db 389 GGGCTACGCTCTGGCTCTCTATGTTCTCCGCGAGCCCGCGAGGTGTGTGGAGNAGTG 448
QY 408 uLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCysLeuLeuArgG 428
Db 449 CTTCACTGTCTGTCTCCAGAACCCCTTGTCTCTGATATATAGGGGTGCTCTATCCGG 508
QY 428 nAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg 448
Db 509 CA--CCCAAGAACCTTCTTCTGTTGGCTCAGCCAGCCACCCACAGCAGTCTTCGAGA 566
QY 448 uLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeu 468
Db 567 GCTCTGGCAACCTGCTGGGATGGGGGTGCTGACCTAGTAGTGGGTGGCAGTCCCTCAC 626
QY 468 rSerCysCysAlaProArgProLysGlyLysSerAsnLeuLeuValValArgArgGlyCy 488
Db 627 TTCCTGTGTATCCCGCAGACCCCAAGAAAGTCCAACTGATCTGTGTCCAGAGAGTCA 686
QY 488 sAsnProAlaProAlaProGlyCysSerPro---SerCysCysAlaLeuThrGlnLeuSe 507
Db 687 CAGCCACCCCATCATCTCTGTTTTCAGCCCAATCCCAATACCACTGAGTCAGTACATGAC 746
QY 507 rPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGlyHisGlySerPh 527
Db 747 ATTTTCAAGATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 806
QY 527 eThrLysIlePheArgGlySerArgGluValValAspGlyGluThrHisAspSerG 547
Db 807 CACCAGATTTTACCGGGGCTGTGCGCATGAGTGTGTGAGTGGAGGCGGCCGAAAGACAGA 866

```

```

QY 547 uValLeuLeuValMetSerArgHisArgAsnCysMetCysSerPheLeuGluAl 567
Db 867 GGTGCTGCTGAAGTCTAGTATGATGCAAGCAAGAACTGATGAGTCAATTCCTGGAAAGC 926

QY 567 aAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMe 587
Db 927 AGGAGCTTGTATGAGCAAGTGTCTACCGGCACTCTGCTGCTCCACGGGCTGTGAT 986

QY 587 tAlaGlyAspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLe 607
Db 987 GGCTGGAGA-MAGACCATGGK-GAGGAATTGTACA-CTGGGGCCATARAMATGTATCT 1043

QY 607 uArgLys 609
Db 1044 GCGAAAA 1050

RESULT 7
BI557690 826 bp mRNA linear EST 05-SEP-2001
DEFINITION 603237050F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5289949 5',
mRNA sequence.
ACCESSION BI557690
VERSION BI557690.1 GI:15445004
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 826)
NTH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11733 row: d column: 14
High quality sequence stop: 760.
FEATURES
source
Location/Qualifiers
1..826
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5289949"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam3"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
```

```

BASE COUNT 147 a 252 c 240 g 187 t
ORIGIN
Alignment Scores:
Pred. No.: 6,44e-95 Length: 826
Score: 1192.00 Matches: 249
Percent Similarity: 95.45% Conservative: 3
Best local Similarity: 94.32% Mismatches: 9
Query Match: 20.34% Indels: 7
DB: 12 Gaps: 1
US-09-397-967a-16 (1-1099) x BI557690 (1-826)
```

```

QY 1 MetAlaProProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
Db 37 ATGGCACTCCAGTGGAGAGACACCTCTGATCCCTCAGCGCTTTCAGCCTCTCATCC 96

QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
Db 97 TCAGAGCAGGAGCCCTGTCATGTCTCTCTCCCGGGGACCTGGGCTCCCGACGGA 156

QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
Db 157 TTGTCAATTC- TTGGGGAGACTCTTGCTGAGAGATTATGTGGCAGCTGCCAGGCC 215

QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db 216 TGTGGCATCTGCTGTTTATCATTCGCTCTGGCCACTGAGGACTTCTCTTTC 275

QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db 276 TGGTTTCCCCCAGCCACATCTTCTGCATAGAGAGCGTGGACACTCAAGTCTTGGTCTAC 335

QY 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120
Db 336 AGGCTACGCTTTTATTTCCCTGACTGTTGGGCTGGAGACATGTACCGCTT- GGGCTG 394

QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
Db 395 GCGAAAGATTG- ACCAGTGGCATCTTGTACTTACATGTTTGTAGAACATCTCTTTGCTCAG 453

QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
Db 454 CACCGCAGTGACCTGTTGAGTGGGCGCTCCCGTGGGCTTAGCATGAAGAGCAGGGA 513

QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
Db 514 GAGTTCCTGAGCCCTGGCGCTGCTGGACTTGGCCAGATGGCTCTGAGCAGGCCCGAGCC 573

QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
Db 574 CCAGAGAGCTGCTGAGAGCGTCTAGTTACAAAGCTCTCTCGCGCCCGCAGCTCGCGAT 633

QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgArg-IleArgArgThrValValLeuAl 220
Db 634 GTGATCCAGGCCAGAACTTCTGTACAGCAGCAGGACATCCGCGAGCCTGGTCTTGC 693

QY 220 aLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLe 240
Db 694 GCTGCG-CCGTGTGCTCGCTTGGAGGCCGACCGTACGCGCTCATGGCCAAAGTATATCT 752

QY 240 uAspLeuGluArgLeuHisPro-AlaAlaThrThrGluThrPheArgValGlyLeuProG 260
Db 753 GGACTTGGAAACGGTACATCCAGAGCGGCAACAACGAGA--CTCGCTGTGGGGCTCCCGG 809

QY 250 lValaGln 262
Db 810 GCGCCAGG 817

RESULT 8
AK087255 2222 bp mRNA linear HTC 05-DEC-2002
LOCUS Mus musculus 0 day neonate lung cDNA, RIKEN full-length enriched
DEFINITION library, clone:E030040D20 product:Janus kinase 2, full insert
sequence.
ACCESSION AK087255
VERSION AK087255.1 GI:26104150
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE Carninci, P. and Hayashizaki, Y.
AUTHORS High-efficiency full-length cDNA cloning
TITLE
```


NIH BMAP M S3.3, NIH BMAP M S3.2, NIH BMAP M S3.1, NIH BMAP M S2, NIH BMAP M S1. The subtracted library (NIH BMAP M S4) was constructed as follows: PCR-amplified cDNA inserts from NIH BMAP M S3.3, NIH BMAP M S3.2, and NIH BMAP M S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH BMAP M S3.3, NIH BMAP M S3.2, and NIH BMAP M S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH BMAP M S4 library. This procedure has been previously described (Bonaldi, Lennon and Soares, Genome Research 6:791-806, 1996).

BASE COUNT 132 a 211 c 215 g 138 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 1,62e-90 Length: 698
Score: 1141.00 Matches: 225
Percent Similarity: 96.15% Conservative: 0
Best Local Similarity: 96.15% Mismatches: 6
Query Match: 19,47% Indels: 4
DB: 12 Gaps: 2

US-09-397-967A-16 (1-1099) x BM935182 (1-698)

QY 792 SerProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAla 811
Db 3 AGTCCTCGAGATGAGCTGTGC---GGTGGCGCCAGCTCTATGCTGCCAGGACCCCGC 59
QY 812 IlePheGluGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySer 831
Db 60 ATATTCGAGAGAGACACCTTAACTACATCTCTTTGCTGGCAAGGCACTTTGGCAGC 119
QY 832 ValGluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLys 851
Db 120 GTGGAGCTGTCCGCTATGACCCCTGGGGGACAAATACGGGACCTGTGGCAGTGAA 179
QY 852 GlnLeuGlnHisSerValProAspGlnGlnArgAspPheGlnArgGluLeuGlnLeu 871
Db 180 CAGCTACACACAGCGGGCCAGACACGACGAGGAGACTTCAGCGGAGATTACAGATCCT 239
QY 872 LysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArg 891
Db 240 AAGGCTCTGCACAGCAGCTTATCTGTCAGTACCGGGAGTACGATATGGCCAGTCGC 299
QY 892 GlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGln 911
Db 300 CAGAGCTCGGTTGGTGATGGATGATACCTGCCAGCGGCTGCTGGAGACTTCCTGCAG 359
QY 912 ArgHisArgGly---LeuHisThrAspArgLeuLeuLeuPheAlaTrpGlnLeuCysLys 930
Db 360 CCCCATCGCGCGCGCTGCACACCGACCCCTACTGCTGCTTGGCAGATCTGCAAG 419
QY 931 GlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIle 950
Db 420 GGCATGGAGTACTTGGTGCGCGCGCTGCTACACCGTACCTGCTGCGCGCAACATC 479
QY 951 LeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuPro 970
Db 480 TTGGTGGAGCAGCGAGCTCATGTGAAGATCGGGACTTCGGCTCGTAAGCTGCTGCC 539
QY 971 LeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerProIle-PheTyrTrpAl 990
Db 540 CTGGGAAGAGACTACTACGTGGTCCGCGAGCTGCGCAAGCCCATCTTNTTGGTATGC 599
QY 990 aProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyVa 1010
Db 600 CCCGAGTCCCTATCTGCAC-ATCTTCTCNCGCCAATCTGACGTGTGGAGCTTCGGAGT 658
QY 1010 lValLeuTyrGluLeuPheThrTyrCysAspLysSerCys 1023

Db 659 GGTTGTTGTACGAGCTCTTACCTACTGCGACAGAGCTGC 698
RESULT 10
BF150250

LOCUS

DEFINITION

BF150250 675 bp mRNA linear EST 29-DEC-2000
uy83e09.y1 NCI CGAP Mam5 Mus musculus cDNA clone IMAGE:3666184 5'
similar to TR:P97423 P97423 JANUS KINASE 3 ; mRNA sequence.

ACCESSION

BF150250

VERSION

BF150250.1 GI:11031645

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Other_ESTs: uy83e09.x1

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

image.llnl.gov/image/html/iresources.shtml

MGI:1426952

Seq primer: -40RP from Gibco

High quality sequence stop: 390.

Location/Qualifiers

1. .675

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:3666184"

/tissue_type="tumor, gross tissue"

/dev_stage="7 months"

/lab_host="DH10B"

/clone_lib="NCI_CGAP_Mams"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigators

providing samples: Lothar Hennighausen/Robin Humphreys,

NIH"

BASE COUNT 138 a 215 c 197 g 123 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 6,41e-90 Length: 675

Score: 1134.00 Matches: 217

Percent Similarity: 96.02% Conservative: 0

Best Local Similarity: 96.02% Mismatches: 7

Query Match: 19,35% Indels: 2

DB: 10 Gaps: 2

US-09-397-967A-16 (1-1099) x BF150250 (1-675)

QY 705 GlyPheGlyAlaThrTrpGluValPheGlnArgGlyProAlaHisIleThrSerLeu 724

Db 1 GGCTTTGGAGCCACCACCTGGGAGTGTTCAGCGGGGACCCGCCACATCACCTCGCTG 60

QY 725 GluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeuProAlaLeuLysTrp 744

Db 61 GAGCCCGCCAAAGCTGAAGTCTTATGAGGACCGAGGACAGCTCCCGCTCTCAATGG 120

QY 745 ThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGlyArgArgProSer 764

Db 121 ACAGAACTGGCGGACTTATCACACAGTCATGGCGTATGATCTCTGCGCGCGCCCTCC 180
 QY 765 PhAArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyrGluLeuLeuSer 784
 Db 181 TTCCGAGCTATCTCTCAGAGACCTCAACGGCTCATTTACATCAGATTACGAGCTCTCTCA 240
 QY 785 AspProThrProGlyIleProSerProArgAspGluLeuCysValAlaGlyAlaGlnLeu 804
 Db 241 GACCCACACTGGCATCCCGAGTCCTCGAGATGAGCTGTGC-----GGTGGCGCCACCTC 297
 QY 805 TyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuIleTyrIleSerLeuLeu 824
 Db 298 TATGCTCTGCGAGACCCCGCATATTCGAGGAGACACCTTAACTAGTACATCTCTTTGCTG 357
 QY 825 GlyIleGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeuGlyIleAspAsnThr 844
 Db 358 GGCAAGGCGCACTTTGGCAGCGTGGAGCTGTGCGCGCTATGACCCCTGGGGGACAATACG 417
 QY 845 GlyProLeuValAlaValIleGlnLeuGlnHisSerValProAspGlnGlnArgAspPhe 864
 Db 418 GGACCCCTGTGGCAGTGGACACTACAGCAGCGGGCCAGACCCAGCAGGAGGACTTC 477
 QY 865 GlnArgGluIleGlnIleLeuIleAlaLeuHisSerAspPheIleValIleTyrArgGly 884
 Db 478 CAGCGGAGATTCAGATCCTTAAGCTCTGACAGCGACTTCATCGTCAAGTACCGGCGGA 537
 QY 885 ValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGly 904
 Db 538 GTACGCTATGGCCGANGTCGCGAGCGCTGGCGTGGTGTGATGGAGTACCTGCCAGCGGC 597
 QY 905 CysLeuArgAspLeuGlnArgHisArgGly--LeuHisThrAspArgLeuLeuLeu 923
 Db 598 TGCCTGGAGACTTCTGACGGCCATCGCGCGCTGACCGGCTGACCGGCTACTGCTGCTG 657
 QY 924 PheAlaTrpGlnIleCys 929
 Db 658 TTCGCTGGCAGATCTGC 675

RESULT 11
 BY730320
 LOCUS
 DEFINITION BY730320 RIKEN full-length enriched, 13 days embryo heart Mus
 musculus cDNA clone D330022C11 5', mRNA sequence.
 BY730320.1 GI:27143447
 EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 677)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,
 Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
 Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,
 Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S.,
 Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani,
 L.E., Cousins, S., Dalia, E., Dragani, I.A., Fletcher, C.F., Forrest,
 A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
 Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,
 Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,
 King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons,
 P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
 H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G.,
 Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.D., Qi, D.,
 Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring,
 B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou,
 M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale,
 R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,
 Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa,
 M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
 Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,
 M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.

, Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
 Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
 K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
 E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs

JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Nature 420, 563-573 (2002)
 12466851
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda,
 S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K.,
 Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno,
 H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K.,
 Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H.,
 Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y.,
 Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct
 Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.

FEATURES
 source

Location/Qualifiers
 1..677
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="D330022C11"
 /tissue_type="heart"
 /dev_stage="13 days embryo"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 13 days embryo
 heart"
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5',
 GAGAGAGAGCGCGCGCACTCGAGTCTTTTCTTTTCTTTT 3'], cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. Second strand cDNA was prepared with the
 primer adapter of sequence [5'
 GAGAGAGAGATTCGAGTCTTAATTAATATCCCCCCCCCC 3']. cDNA
 was cleaved with BamHI and XhoI. Vector: a modified
 pBluescript KS(+) after bulk excision from lambda FLIC 1."

BASE COUNT
 ORIGIN

117 a 203 c 188 g 169 t


```

Db      603  GACCCAGGAGACAGAGGCTCCAGCCCTCTCGAGATTCAGAGATCGTAGACAT 662
Qy      297  lserlIeAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValTh 317
Db      663  TAGCATCAAGCAGGCCCCGCGCTTGCGCGCGCGAGAGACACGCGCTGCTCACTGTTAC 722
Qy      317  rArqMeTAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPh 337
Db      723  CAGACACAGACACACAGATTTTAGAGGCGGAGTTCCTCCAGGCTGCCAGGCTGTGCTG 782
Qy      337  eValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLy 357
Db      783  CGTGGCGCTCGTGACGCGCTACTTCGCGTGACACGAGCTCCAGCAGACTTCTTCTGCG 842
Qy      357  sGluValAlaProProArgLeuLeuGluGluGluAlaAspValCysHisGlyProIleTh 377
Db      843  GAGGTTGCAACCGCGCGAGCTGTG-NAGGAAAGTGCCNANCAGTGCCA-GGNCCTCAC 900
Qy      377  rLeuAsp 379
Db      901  TCTGGAC 907

```

RESULT 13

```

BY746452
LOCUS      BY746452          655 bp      mRNA      linear      EST 17-DEC-2002
DEFINITION BY746452 RIKEN full-length enriched, 2 days neonate thymus thymic
cells (NOD) Mus musculus cDNA clone E430014A10 5', mRNA sequence.
VERSION     BY746452.1 GI:27174142
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus

```

```

REFERENCE   1 (bases 1 to 655)
AUTHORS     Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaado,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H.,
            Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C.,
            Gojobori,T., Baidarelli,R., Hill,D.P., Bulc,C., Hume,D.A.,
            Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S.,
            Beisel,K.W., Blake,J.A., Bradt,D., Brusick,V., Chocthia,C., Corbani
            L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest
            A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A.,
            Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J.,
            Jarvis,E.D., Kanaia,A., Kawai,H., Kawasawa,Y., Kedzierski,R.M.,
            King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons
            P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki
            H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G.,
            Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D.,
            Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring
            M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale
            R.D., Tomita,M., Verardo,R., Wagner,L., Wahlstedt,C., Wang,Y.,
            Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa
            M., Yang,I., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A.,
            Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura
            M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii
            Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata
            K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander
            E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

```

```

Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs

```

```

Nature 420, 563-573 (2002)

```

```

JOURNAL     22354683
MEDLINE     12466851
PUBMED

```

```

COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center (GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-2 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222

```

```

Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda
S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F., Imotani,K.,
Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo,S., Konno
H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K.,
Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N., Sano,H.,
Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y.,
Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct
Submission

```

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Reissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC Building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

```

Location/Qualifiers
1..655
/mol_type="mRNA"
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="E430014A10"
/tissue_type="thymus"
/cell_type="thymic cells"
/clone_lib="RIKEN full-length enriched, 2 days neonate thymus thymic cells (NOD)"
BASE COUNT 111 a 197 c 180 g 167 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 6,42e-81 Length: 655
Score: 1032.00 Matches: 195
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.61% Indels: 0
DB: 14 Gaps: 0
US-09-397-967a-16 (1-1099) x BY746452 (1-655)

```

```

Qy      1  MetalProProSerGluGluThrProIleuLeuProGlnArgSerCysSerLeuSerSer 20
Db      63  ATGCACCTTCAAGTGAGGACACCTCTGTATCTCTCAGCGCTCTTGAGGCTCTCATACC 122
Qy      21  SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
Db      123 TCAGAGGAGGAGGAGCGCTGCATGTCTCTCTCTCCCGGGAGCGCTGGCGCTCCACGCGA 182
Qy      41  LeuSerPheSerPheGlyAspTyrIleuAlaGluAspLeuCysValArgAlaAlaLysAla 60
Db      183 TTGTCACTCTCTTTTGGGGACTACTTGGCTGAGGATTTATGTGCGAGCTGCCAAGGCC 242
Qy      61  CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db      243 TGTGGCATCTCGCTGTTTATCTTCGCTTTGCTTGGCCACTGAGACTTCTCTTTC 302

```



```
QY 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db 303 TGGTTTCCCCCAAGCCACACTTCTGCATAGAGGAGCTGGACACTCAAGTCTTGGTCTAC 362
QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
Db 363 AGGCTACGCTTTATTTCCCTGACTGGTTTGGGCTGGAGACATGTCACCGCTTTGGGCTG 422
QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
Db 423 CGCAAGATTGACCATGGCCATCTTACATGCTTGAAGATCTCTTTGCTCAG 482
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
Db 483 CACCGCAGTGACCTGTGTAGTGGCGCTCCCGGTGGCCCTTAGCATGAAGACAGGGA 542
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
Db 543 GAGTTCTGAGCGCTGGCCGTGTGGACTTGGCCAGATGCTGTCGAGCAGGCCAGCGC 602
QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuPro 195
Db 603 CCAGGAGAGCTGCTGAAGACGGTCAGTTACAAAGCCTGTCTGCCG 647

RESULT 14
BG870513 907 bp mRNA linear EST 29-MAY-2001
LOCUS 602791432F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4922742 5',
DEFINITION mRNA sequence.
ACCESSION BG870513
VERSION BG870513.1 GI:14221053
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 907)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10842 row: p column: 07
High quality sequence stop: 742.
FEATURES
Location/Qualifiers
1..907
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4922742"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 166 a 279 c 275 g 185 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 1,16e-80 Length: 907
Score: 1031.50 Matches: 216
Percent Similarity: 88.93% Conservative: 1
Best Local Similarity: 88.52% Mismatches: 23
Query Match: 17.60% Indels: 7
```

```
DB: 12 Gaps: 3
US-09-397-967A-16 (1-1099) x BG870513 (1-907)
QY 858 ProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAsp 877
Db 39 CCAGACCAGCAGAGGGAGACTCCAGCGGGAGATTCCAGATCTCTTAAGGCTCTGCACAGCGAC 98
QY 878 PheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuVal 897
Db 99 TTCATCGTCAAGTACACGGGGAGTCAGCTATGGGCCAGGTCGCCAGAGCCCTCGCGTTGGTG 158
QY 898 MetGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGly---Leu 916
Db 159 ATGGAGTACTGCCAGAGCGGCTGCCAGAGACTTCTGCGAGCGCCATCGCGCGCGCTG 218
QY 917 HisThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGly 936
Db 219 CACACCGACCGCCTACTGCTGTTGCTTGGCAGATCTGCAAGGGCATGGAGTACCTGGGT 278
QY 937 AlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAla 956
Db 279 GCGCGCCGCTGAGTACACCGAGACCTGGTGGCGGCAACATCTTGTGGAGAGCGAGGCT 338
QY 957 HisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyr 976
Db 339 CATGTGAAGATCGCGACTTCGGCCTCGCTAAGCTGCTGCCCTGGGAAGAGACTACTAC 398
QY 977 ValValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLeuSerAsp 996
Db 399 GTGGTCCGCGAGCCTGGCCAAAGCCCATCTTTGGTATGCCCGCGAGNTCTATCTGAC 458
QY 997 AsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPhe 1016
Db 459 AACATCTTCTCCGCGCAATCTGACGTGTGGAGCTTCGGAGTGGTGTGTACGAGCTCTTC 518
QY 1017 ThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGlu 1036
Db 519 AC-TACTGGCAGCAGAGCTGCAGNCCATCCGCTGAGTTCCTGAGCATGATGGGCGCTGAG 577
QY 1037 ArgGluGlyProProLeuCysArgLeuLeuLeuLeuAlaGluGlyArgArgLeuPro 1056
Db 578 CGTGAAGGACCCCGTTCTTTCGCTC-CTGGAGCTGCTGCAGAGGCGCGAGC-CTCCCA 635
QY 1057 ProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGlu 1076
Db 636 CATCTTTCCACTGCC---CACGAGGTTTCAGAGCTCATGAGCTGTGCTGGCGCCGACG 692
QY 1077 ProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGly 1096
Db 693 CCG---ACGACGGGAGCGCTTCGGACCTTCAGCCCTGAGCCCCCAGCTGCACCGCTGTGGAGTGA 749
QY 1097 -ArgProGly 1099
Db 750 AAGACCCGGG 759

RESULT 15
BG873355
LOCUS 602794328F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4925598 5',
DEFINITION mRNA sequence.
ACCESSION BG873355
VERSION BG873355.1 GI:14223895
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 720)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
```

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10850 row: 9 column: 07
High quality sequence stop: 709.
Location/Qualifiers
1. .720
source
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4925598"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 Kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."
BASE COUNT 123 a 239 c 229 g 129 t
ORIGIN

Alignment Scores:
Pred. No.: 1.51e-80 Length: 720
Score: 1028.50 Matches: 201
Percent Similarity: 95.71% Conservative: 0
Best Local Similarity: 95.71% Mismatches: 8
Query Match: 17.55% Indels: 3
DB: 12 Gaps: 1

US-09-397-967A-16 (1-1099) x BG873355 (1-720)

Qy 891 ArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeu 910
Db 1 CGCCAGAGCCCTGGGTTGGTGTGATGGAGTACTGCCAGCGGCTGCTCGGAGACTTCCTG 60
Qy 911 GlnArgHisArgGly--LeuHisThrAspArgLeuLeuPheAlaTyrGlnIleCys 929
Db 61 CAGCGCCATCGCGCGCGCTGCACACCGCGCTACTGCTGTTGCTTGGCAGATCTGC 120
Qy 930 LysGlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsn 949
Db 121 AAGGGCATGGAGTACCTGGGTGGCGCGCTGCGGTACACCGTACCGTGGCTGGCGGCAAC 180
Qy 950 IleLeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeu 969
Db 181 ATCTTGGTGGAGAGGAGGCTCATGTGAAGATCGGGACTTCGGCTCGCTAAGCTGCTG 240
Qy 970 ProLeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerProIlePheTyrPyr 989
Db 241 CCCCTGGGAAAGGACTACTAGTGTGCGGAGCTGGCCAAAGCCCATCTTTTGGTAT 300
Qy 990 AlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTyrSerPheGly 1009
Db 301 GCCCGGGAGTCCCTATCTGACAACATCTTCCCGCCCAATCTGACGTGTGGAGCTTCGGA 360
Qy 1010 ValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPhe 1029
Db 361 GTGGTGTGTACGAGCTCTTACCTACTTCGACAGAGCTGCAGCCCATCCGCTGAGTTC 420
Qy 1030 LeuArgMetMetGlyProGluArgGluGlyProProLeuCysArgLeuLeuLeuLeu 1049
Db 421 CTGGCATGATGGGGCCTGAGGCTGAGGAGACCCCGCTCTGCGGCTCTCTGGAGCTGCTG 480
Qy 1050 AlaGluGlyArgArgLeuProProProProThrCysProThrGluValGlnGluLeuMet 1069
Db 481 GCAGAGGGCGGACGGCTCCCACTCCACCACTCCACCTCCCGCCACCGAGGTTTCAGGAGCTCATG 540
Qy 1070 GlnLeuCysTyrAlaProGluProHisAspArgProAlaPheAlaThrLeuSerProGln 1089

Db 541 CAGCTGTGTGGGCGCCAGCCCGCA-GACCGGCGAGCTTCGG-ACCCTGAGCCCCCAG 598
Qy 1090 LeuAspProLeuTyrArgGlyArgProGly 1099
Db 599 CTGGACGCGCTGTGGCGTGAAGACCCGGA 628

Search completed: February 4, 2004, 08:14:50
Job time : 6493.54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2004, 01:30:36 ; Search time 172.373 Seconds
(without alignments)
2814.136 Million cell updates/sec

Title: US-09-397-967A-16

Perfect score: 5860

Sequence: 1 MAPSEETPLPQRCSLSS.....RPAFATLSPQLDPLWRGREG 1099

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09397967/runat_03022004.175632.28596/app_query.fasta_1.1486
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINWATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pot -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09397967 @CGN 1 1 81 @runat_03022004.175632.28596 -NCPU=6 -ICPU=3
-NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*

1:	/cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2:	/cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3:	/cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4:	/cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5:	/cgn2_6/ptodata/1/ina/6C.COMB.seq:*
6:	/cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5223	89.1	4016	5	PCT-US95-08354A-1
2	4597	78.4	3620	4	US-09-016-434-297
3	4418	75.4	3807	2	US-08-357-598-1
4	4418	75.4	3807	2	US-08-003-289-1
5	4418	75.4	3807	5	PCT-US95-16435-1
6	2629	44.9	3435	3	US-09-046-158A-21
7	2621	44.7	4482	2	US-08-567-508C-1
8	2621	44.7	4482	3	US-09-196-480-1
9	2621	44.7	5117	4	US-09-972-800A-15
10	2619.5	44.7	3629	1	US-08-097-997A-8
11	2619.5	44.7	3629	3	US-08-665-574C-8
12	2619.5	44.7	3629	3	US-08-946-994-8

13	2594	44.3	4078	4	US-09-016-434-297	Sequence 297, Appl
14	2360	40.3	3495	1	US-08-446-038B-2	Sequence 2, Appli
15	2360	40.3	3495	1	US-08-446-010B-2	Sequence 2, Appli
16	2360	40.3	3495	1	US-08-805-445-2	Sequence 2, Appli
17	2360	40.3	3495	2	US-08-064-067D-2	Sequence 2, Appli
18	2360	40.3	3495	2	US-09-066-208-2	Sequence 2, Appli
19	2360	40.3	3495	4	US-08-980-080-3	Sequence 3, Appli
20	1904	32.5	3429	1	US-08-097-997A-10	Sequence 10, Appl
21	1904	32.5	3429	3	US-08-665-574C-10	Sequence 10, Appl
22	1904	32.5	3429	3	US-08-946-994-10	Sequence 10, Appl
23	1904	32.5	4234	1	US-08-446-038B-1	Sequence 1, Appli
24	1904	32.5	4234	1	US-08-446-010B-1	Sequence 1, Appli
25	1904	32.5	4234	1	US-08-805-445-1	Sequence 1, Appli
26	1904	32.5	4234	2	US-08-064-067D-1	Sequence 1, Appli
27	1904	32.5	4234	2	US-09-066-208-1	Sequence 1, Appli
28	1817.5	31.0	4080	4	US-09-016-434-1353	Sequence 1353, Ap
29	1817.5	31.0	4176	4	US-09-972-800A-17	Sequence 17, Appl
30	1815.5	31.0	3561	1	US-08-097-997A-12	Sequence 12, Appl
31	1815.5	31.0	3561	3	US-08-665-574C-12	Sequence 12, Appl
32	1815.5	31.0	3561	3	US-08-946-994-12	Sequence 12, Appl
33	482	8.2	3845	2	US-08-220-240A-4	Sequence 4, Appli
34	475	8.1	4530	1	US-08-229-515A-9	Sequence 9, Appli
35	475	8.1	4530	4	US-09-167-322-4	Sequence 4, Appli
36	475	8.1	4530	4	US-09-527-487-1	Sequence 1, Appli
37	475	8.1	4530	4	US-09-877-177A-11	Sequence 11, Appl
38	474	8.1	3768	2	US-08-625-101-1	Sequence 1, Appli
39	474	8.1	3768	2	US-08-356-786-1	Sequence 1, Appli
40	474	8.1	4473	2	US-09-048-804-1	Sequence 1, Appli
41	474	8.1	4473	3	US-09-056-105-26	Sequence 26, Appl
42	474	8.1	4473	3	US-08-475-035-3	Sequence 3, Appli
43	466.5	8.0	5532	2	US-09-676-610B-17	Sequence 17, Appl
44	466.5	8.0	5532	4	US-09-632-580A-3	Sequence 3, Appli
45	465.5	7.9	5484	3		

ALIGNMENTS

RESULT 1
PCT-US95-08354A-1
; Sequence 1, Application PC/TUS9508354A
; GENERAL INFORMATION:
; APPLICANT: Temple University - Of The
; APPLICANT: Commonwealth System of Higher Education
; TITLE OF INVENTION: JAK3 PROTEIN TYROSINE
; TITLE OF INVENTION: KINASE AND DNA ENCODING THE SAME
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorigna
; ADDRESSEE: & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08354A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/272,368
; FILING DATE: 8 July 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-203 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549

QY 421 TyrIysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSer 440
Db 1921 TACAAGGGCTGCTCATCCGCCAGGACCCAGCGGGCTTCTCCCTGGTTGC-CTCAGC 1979
QY 441 GlnProHisArgSerLeuArgGluLeuLeuAlaAalaCysTrpAsnSerGlyLeuArgVal 460
Db 1980 CAGCCCCACAGAAAGCTCGCGGAGCTGCTTGCAGC-TGCTGGAAATTC-GGGCTGCGAGTA 2037
QY 461 AspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGlyLysSerAsn 480
Db 2038 GACGGTGTCTGCTGAACCTACATCTCTGCTGGCTCCAGACCCAGAAAGTCCAAT 2097
QY 481 LeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCys 500
Db 2098 TTGATCGTGTGCGAAGGGGCTGCACCCCGCGCTGCCCTGGCTGCTCCCGTCTGTC 2157
QY 501 CysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGlu 520
Db 2158 TGTGGCTGACACAGCTGAGCTTCCACAAATTCACAGCGACAGCCTG--GGACACGAG 2214
QY 521 AsnLeuGlyHisGlySerPheThrIysIlePheArgGlySerArgArgGluValValAsp 540
Db 2215 AACCTGGGTTCACGGTCTTTTACCAAGATCTTCCGTGGCGCAGCGGGAGGTGCGGAT 2274
QY 541 GlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCys 560
Db 2275 GGTGAGACACATGACTCGGAAGTCTCTCTGAAGTTCATGACTCCAGACATCGGAATGC 2334
QY 561 MetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuVal 580
Db 2335 ATGGAGTCTTTTCTGGAGCCGCAAGCTTGATGAGCCAAATTCCTACCGCACCTGGTG 2394
QY 581 LeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTyrLeu 600
Db 2395 TTACTGCACGGCTGTGCATGGCTGGAGACACATCATGTGTCAGAAATTTGTGTATCTA 2454
QY 601 GlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLysLeu 620
Db 2455 GGAGCAATTGATGACTGCTGCGAAGCGTGCCACCTGGTGTGACGCCAGCTGGAACTG 2514
QY 621 GlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHis 640
Db 2515 CAGGTGACCAAGCAGCTGGCATATGCCCTTAATCTTGGAGGACAAAGGCGCTTCTCTAC 2574
QY 641 GlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyValAspGlyAsnProPro 660
Db 2575 GGCAAGGTCTCAGCACGGAAGTGTCTCTGGCTCGTGAGGGGGGTGATGGGAATCCACT 2634
QY 661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680
Db 2635 TTCAATTAAGCTGAGTGATCTCTGTGCTCAGTCCCACTGTGTGAGCTGGAATGCTCACC 2694
QY 681 AspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 700
Db 2695 GACAGAAATACCCTGGGTGGCCCCCGGAATGCTCCAGGAGGCTCAGACACTGTCTGTGGAG 2754
QY 701 AlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHis 720
Db 2755 GCTGACAAGTGGGCTTTGGAGCCACCACTGGGAGGTGTTACGGGGGGAGCCCGCCAC 2814
QY 721 IleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeuPro 740
Db 2815 ATCACTCGCTGGAGCCCGCCCAAAAGCTGAAGTTCTATGAGGACCCAGGACAGCTGCC 2874
QY 741 AlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760
Db 2875 GCTCTCAATGGACAGAACTGGCGGACTTATCACAGTGCATGGCGGTATGATCTCTGGC 2934
QY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780
Db 2935 CGGCGCCCTCTTCGAGCTATCTCTCAGAGACCTCAACGCCCTCATATCATCATGATTAC 2994
QY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAla 800

Db 2995 GAGCTCTCTCA-GACCCCAACACCTGGCATCCCGAGTCCTCGAGATGAGCTGTGCGT---- 3049
QY 801 GlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLysTyr 820
Db 3050 GGGCGCCAGCTCTATSCCTGCCAGGACCCCGCCATATC-GAGGAGAGACACCTTAAGTAC 3108
QY 821 IleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeu 840
Db 3109 ATCTCTTTTCTGGCAAGGGCAACTTTGGCAGCGTGGAGCTGTGCGCTATGACCCCTG 3168
QY 841 GlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAspGln 860
Db 3169 ---GACAAATACCGGACCCCTGGTGGCAGTGAACAGCTACAGCACAGCGGGCAGACCA 3225
QY 861 GlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVal 880
Db 3226 CAGAGGAGCTTCCAGCGGAGATTCAGATCTTAAAGCTTCTGCACAGCGACTTCATCGTC 3285
QY 881 LysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTyr 900
Db 3286 AAGTACCGGGAGTCAGCTATGGGCAGTCCAGAGCGCTCGCGCGGCCCTGTCACACCGAC 3345
QY 901 LeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHisThrAsp 919
Db 3346 CTGCCCAGCGGCTGCTGGAGACTTCTGACAGCCCATCTCGCGCGGCCCTGTCACACCGAC 3405
QY 920 ArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArgArg 939
Db 3406 CGCCTACTGCTTTCGCTTGGCAGATCTGCAAGGCGATGGAGTACTCTGGTGGCGCGCGC 3465
QY 940 CysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisValLys 959
Db 3466 TCGGTACACCGGTGACTGCTGGCGCAACATCTTGGTGGAGAGCGAGGCTCATGTGAAG 3525
QY 960 IleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArg 979
Db 3526 ATCGCGGACTTTCGGCTCGTAAAGCTGCTGCCCTTGGGAAAGGACTACTACGTGTCGCG 3585
QY 980 GluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnIlePhe 999
Db 3586 GAGCTGSCCAAAAGCCCATCTTTTGGTATGCCCGGAGTCCCTATCTGACAAATCTTC 3645
QY 1000 SerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyrCys 1019
Db 3646 TCCCGCCAAATCTGAGTGTGAGCTTCGGAGTGTGTGTACGAGCTCTTCACTTACTGTC 3705
QY 1020 AspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGluGly 1039
Db 3706 GACAAGAGCTCAGCCCATCCGCTGAGTTCTGCGCATGATGGGGCTGAGCGTGAAGGA 3765
QY 1040 ProProLeuCysArgLeuLeuLeuAlaGluGlyArgArgLeuProProProPro 1059
Db 3766 CCCCCTGCTGCGGCTCTGAGCTGCTGGCAGAGGGCCAGCCCTCCCACTTCTCC 3825
QY 1060 ThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGluProHisAsp 1079
Db 3826 ACCTGCCCCACCGAGGTTACAGAGCTCATGCACTGTGCGTGGCGGCCAG-CGACACGAC 3884
QY 1080 ArgProAlaPheAlaThrLeuSerProGluLeuAspProLeuTrpArgGlyArgPro 1098
Db 3885 CGGCGACCTTGGCACCTTGGCCCCCAGCTGACGCGCTGTGGCGTGGGAAGACCC 3941

RESULT 2

US-09-016-434-1055
; Sequence 1055, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490

CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
FILING DATE: HEREWITH
FILING DATE: 09/016,434
PRIORITY APPLICATION DATA:
CLASSIFICATION:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1055:
SEQUENCE CHARACTERISTICS:
LENGTH: 3620 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: G1039418
US-09-016-434-1055

Alignment Scores:
Pred. No.: 0 Length: 3620
Score: 4597.00 Matches: 895
Percent Similarity: 86.80% Conservative: 65
Best Local Similarity: 80.92% Mismatches: 123
Query Match: 78.45% Indels: 24
DB: 4 Gaps: 7

US-09-397-967a-16 (1-1099) x US-09-016-434-1055 (1-3620)

QY 1 MetAlaProSerGluGluThrProLeuLeuProGlnArgSerCysSerLeuSerSer 20
Db 96 ATGGACCTTCAAGTGAAGACGCCCTGATCCCTCAGCGTTTCATGAGCGCTTGTGCC 155
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProArgGlyProGlyProGlnArg 40
Db 156 ACGAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 215
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAla 60
Db 216 CTATCTTCTCTTGGGACCACTTGTGCTGAGGACCTGTGCTGAGGCTGCGTGCAGG 275
QY 61 CysGlyLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db 276 AGCGGACATCTGCTGTATACCTCTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 335
QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db 336 TGGTTCCTCCCGAGCCACATCTTCTCCGTGGAGGATGCCAGCACCCCAAGTCTGTGTAC 395
QY 101 ArgLeuArgPheTyrProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
Db 396 AGGATTCGCTTTACTTCCCAATTTGGTTGGCTGGAGAGTGCACCGCTTCGGGCTA 455
QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140

Db 456 CGCAAGGATTGGCCAGTGCTATCTTCAGCTGCCAGTCTTGAGACCTCTTTGCCGAG 515
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMethylGlnGly 160
Db 516 CACCGAGTACCTGTGTAGTGGGCTCTCCCGTGGGCTCAGTCTCAAGGACGAGGT 575
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
Db 576 GAGTGTCTCAGCTGCGCGTGTGAGCTTGGCCGATGGCGGAGAGCAGGAGCCGAGCG 635
QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
Db 636 CCGGAGAGCTCTGAAGACTGTACAGGCTGTCTACCTACCTACCTACCTACCTACCTAC 595
QY 201 ValIleGlnGlyGlnAspPheValThrArgArgIleArgArgThrValValLeuAla 220
Db 696 CTGATCCAGGGCTGAGCTTCGTGACGCGAGCGGTATTCGGAGGACGCTGCGCAGG 755
QY 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLe 240
Db 756 CTGCG-CCGCGTGGCGCTGCGCAGGACAGCGGACTCGCTCATGGCCCAAGTACATCAT 814
QY 240 uAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260
Db 815 GGACCTGGAGCGCTGGATCCAGCGGGCGCGGAGACCTTCCACGTGGGCTCTCCCTGG 874
QY 260 YAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePr 277
Db 875 GGCCTTGGTGGCCACGAGCGGCTGGGCTGTCTCCGCTGGTGGTGGTGGTGGTGGTGG 934
QY 277 oTrpSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAs 296
Db 935 CTGACCCAGGAGAGACAGGAGGTCTCTCCAGCGCTTCTGCGACTTTCCAGAAATCGTAGA 994
QY 296 pValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrVa 316
Db 995 CATTAGCATCAAGCAGGCGCGCGGTGGCGCGGCGGAGAGCACCGCTGGTGCATGT 1054
QY 316 lThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSe 336
Db 1055 TACAGGACAGACACACAGATTTTAGAGGCGGAGTCCCGAGGCTGCCGAGGCTGTGTC 1114
QY 336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCy 356
Db 1115 GTTCGTGGCGCTCGTGGACGCTACTTCCGCGTGACCAACGAGCTCCCGAGCATCTTCT 1174
QY 356 sLysGluValAlaProProArgLeuLeuGluAlaAspValCysHisGlyProIl 376
Db 1175 CAAGGAGGTGGCCCGCGGAGGTGCTGGAGGAGTGGCGGAGTGGCGGAGTGGCGGAGT 1234
QY 376 eThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIl 396
Db 1235 CACTCTGGACTTGGCATCAACAGCTCAAGCTGGGCGCTCAGCTCTGCTGCTGCTGCTGT 1294
QY 396 eLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrPr 416
Db 1295 TCTCCGCGGAGCGCCCGAGGACTTGTACAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1354
QY 416 oLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLe 436
Db 1355 CCTTGGTCTGATTATAAGGCTGCTCATCCGCGGAGCGCCCGAGGAGTGGAGTGGAGT 1414
QY 436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaCysTyrAsnSe 456
Db 1415 GTTGGCTCTAGCGGACCCCGAGGAGTCTTCGAGAGCTCTCGGAGCTCTGCGAACCTG 1474
QY 456 rGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProly 476
Db 1475 GGGGCTGACGCTAGATGGGCTGAGCAGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1534
QY 476 sGluLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCy 496

1535 AGAAAGTCCAACTGATCGTGTCCAGAGGTCACAGCCCAACACATCATCTTGGT 1594
QY 496 sSerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSe 515
Db 1595 TCAGCCCCCAATCCCAATACAGCTGAGTCAGATGACATTTCAAGATCCCTGCTGACAG 1654
QY 515 rLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerAr 535
Db 1655 CTGGAGTGGCATGAGAACCTGGGCGCATGGCTCTTCCACCAAGATTACCGGGGCTGTGC 1714
QY 535 gArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSe 555
Db 1715 CCATGAGTGGTGGATGGGAGGCCCGAAGACAGAGGTGCTCTGAAGTCAATGGATGC 1774
QY 555 rArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575
Db 1775 CAAGCACAAAGAACTGCATGGAGTCATCTCTGGAGCAGGAGCTTGATGAGCCAAAGTGC 1834
QY 575 rTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValG1 595
Db 1835 GTACCGGCATCTCGTCTCTCACGGCTGTGCATGGTGGAGACAGCACCATGGTGCA 1894
QY 595 nGluPheValTyrLeuGluValAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSe 615
Db 1895 GGAATTTGTACACCTGGGGGCCATAGACATATCTGCCAAACCGTGGCCACCTGGTGCC 1954
QY 615 rAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAs 635
Db 1955 AGCCAGCTGGAGCTGCAGGTGGTCAACAGCTGGCTACGCCCTCACTATCTGGAGGA 2014
QY 635 pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyG1 655
Db 2015 CAAGCGCTGCCCATGGCATGGATGTCTCTGCCGGAAGGTGCTCTGGCTCGGGAGGGGC 2074
QY 655 yAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSe 675
Db 2075 TGATGGAGCGCGCCCTTCATCAAGCTGAGTGACCTGGGGTCAGCCCGCTGTGTTAAG 2134
QY 675 rLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaG1 695
Db 2135 CTGGAGATGTCACCGACAGGATCCCTTGGGTGGCCCCCGGAGTGCTCTCCGGAGGGCGCA 2194
QY 695 nThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrTrpGluValPheG1 715
Db 2195 GACACTTAGCTTGAAGCTGACAGTGGGGCTTCGGCGCCACGGTCTGGGAAGTGTTAG 2254
QY 715 nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAs 735
Db 2255 TGGCGTCACCATGCCATCAGTGCCCTGGATCTGCTTAAGAACTCCAAATTTATAGGA 2314
QY 735 pGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMe 755
Db 2315 CCGGCAGCAGCTGCGGCCCCCAAGTGGAGACAGAGTGGCCCTGTGATTCAACAGTGGAT 2374
QY 755 tAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775
Db 2375 GGCCTATGAGCCGCTCCAGAGCGCTCTCTCCGAGCGCTCATTCGTGACCTCAATAGCT 2434
QY 775 uIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 795
Db 2435 CAICTTTCAGACTATGAGTCTCTCTCAGACCCACACCTGGTCCCTGGCACCTCGTGA 2494
QY 795 pGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluG1 815
Db 2495 TGGGCTGTGG---AATGGTGGCCAGCTCTATGCTTGCCAGACCCACGATCTTCGAGGA 2551
QY 815 uArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCy 835
Db 2552 GAGACACCTCAAGTACATCTCACAGCTGGGCAAGGGCAACTTTGGCAGCGTGGAGCTGTG 2611
QY 835 sAtqTyrAspProLeuGlyAsnThrGlyProLeuValAlaValLysGlnLeuGlnHi 855
Db 2612 CCGCTATGCCCGCTAGGCGCAATACAGGTGGCCCTGGTGGCGGTGAACAGCTGCAGCA 2671

QY 855 sSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHi 875
Db 2672 CAGCGGGCCAGACACAGCAGAGGACTTTTCAGCGGGAGATTTCAGATCTCTCAAGCACTGCA 2731
QY 875 sSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuAr 895
Db 2732 CAGTGATTTTCATTGTTCAGGTATCTGTGTGTACGTATGTGGCCGGCGCCAGAGCTGGC 2791
QY 895 gLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgG1 915
Db 2792 GCTGTGTCATGGAGTACCTGCCAGCGGCTGCTTGGCGACTTCCTGCAGCGGCACCGCGC 2851
QY 915 y---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTy 934
Db 2852 GCGCTTCGATGCCAGCGGCTCTCTCTCTATTTCTCGCAGATCTGCAAGGGCATGGAGTA 2911
QY 934 rLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSe 954
Db 2912 CTGGGCTCCCGCGCTGGTGCACCGACCTGGCGCCCGCAACATCTCTGTGGAGAG 2971
QY 954 rGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAs 974
Db 2972 CGAGGCACACGTCAGATCGCTGACTTCGGCTTAGCTAGCTGCTGCCCTTCACAAAGA 3031
QY 974 pTyrTyrValValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLe 994
Db 3032 CTACTACGTGGTCCGAGCGAGCGCAGAGCCCATTTTCTGGTATGCCCGGAATCCCT 3091
QY 994 sSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrG1 1014
Db 3092 CTCGGACAACTTCTCTCGCCAGTCAGACGTCTGGAGCTTCGGGGTCTGCTGTACGA 3151
QY 1014 uLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetG1 1034
Db 3152 GCTCTTCACTACTCGCACAAAGCTGACAGCCCTCGCGCGCTCTTGAACCTGCTGGAGGGCCAGAG 3211
QY 1034 yProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgAr 1054
Db 3212 ATGTAGCGGGATGTCCCCGCCCTCTGCGGCTCTTGAACCTGCTGGAGGGCCAGAG 3271
QY 1054 gLeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAl 1074
Db 3272 CTGCGCGGCTCTGCTGCTGCCCTGTGAGGTG----- 3305
QY 1074 aProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTr 1094
Db 3306 -----AGCGCGCGAGGGCTAGCTCAGTTCCAGTCTGTAGAT-----TG 3346
QY 1094 pArgGlyArgProGly 1099
Db 3347 GGCGGGGCTCTCGGGC 3362

RESULT 3

US-08-357-598-1

; Sequence 1, Application US/08357598

; Patent No. 5705625

; GENERAL INFORMATION:

; APPLICANT: Civin, Curt I.

; APPLICANT: Small, Donald

; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: CA

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/357,598

FILING DATE: 15-DEC-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Haile, Lisa A.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 07265/033001

TELEPHONE: 619/678-5070

TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3807 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

US-08-357-598-1

Alignment Scores:
 Pred. No.: 0
 Score: 4418.00
 Percent Similarity: 86.94%
 Best Local Similarity: 81.16%
 Query Match: 75.39%
 DB: 12
 Gaps: 6

US-09-397-967a-16 (1-1099) x US-08-357-598-1 (1-3807)

QY 1 MetAlaProSerGluThrProLeuLeuProGlnArgSerCysSerLeuSer 20
 Db 168 ATGGCACTCCAAAGTGAAGACGCGCCCTGATCCCTCAGCGTTCATGACGCCCTCTTGTC 227
 QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProArgGlyProGlyProGlnArg 40
 Db 228 ACGGAGGCTGGTGGCCCTGCATGTGCTGCTGCGCTGGGGCCCGGGCCCGCCAGCGC 287
 QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
 Db 288 CTATCTTCTCTTTGGGACCACTTGGCTGAGGACCTGTGCGGTGACAGGCTGCCAAGGCC 347
 QY 61 CysGlyLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
 Db 348 AGCGCATCTCGCTGTGTACACTCCCTCTTCTGCTGGCCAGGAGACCTGTCTCTGC 407
 QY 81 TrpPheProProSerHisLeuPheCysIleGluAspValAspThr-GlnValLeuValTy 100
 Db 408 TGGTT-CCCCGAGCCACATCTTCTCGTGAGGATGCCAGCACCCCAAGTCTCTGTGA 466
 QY 100 rArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLe 120
 Db 467 CAGGATTCGCTTTTACTTCCCAATTTGTTGGGTGGAGAAAGTGCACCGCTTCGGGCT 526
 QY 120 uArgLysAspLeuThrSerAlaLeuLeuAspLeuHisValLeuGluHisLeuPheAlaGl 140
 Db 527 ACGCAAGGATTTGGCCAGTGTATCTTCTGACCTGCTGCGGCTGAGGACCTTTTGCCCA 586
 QY 140 nHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGl 160
 Db 587 GCACCGAGTACCTGTGTAGTGGCGCTCTCCCGTGGCTCAGTCTCAAGGAGCAGG 646
 QY 160 yGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnAr 180
 Db 647 TGAGTGTCTCAGCCTGGCGGTGTGAGCTGCGCCGAGTGGCGGAGAGCAGGCCAGCG 706
 QY 180 gProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAs 200
 Db 707 GCGGGAGAGCTGTCTGAAGACGTGTGAGCTCAAGGCCCTGCTTACCCCAAGCCTCGCGA 766
 QY 200 pValileGlnGlyGlnAsnPheValThrArgArgIleArgThrValValLeuAl 220

Db 767 CCGTATCCAGGCGCTGAGCTTCGTACCGGGGAGGGCTATTTCGAGGACGGTGGAGAGCCC 826
 QY 220 aLeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleL 240
 Db 827 CCGTGG-CCGGTGGCCGCTGCCAGGACGACCGGCACTGCTCATGGCAAGTATCATCA 885
 QY 240 euAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260
 Db 886 TGGACCTGGAGCGGCTGGATCCAGCCGGGGCCCGGAGACCTTCCAGTGGGCGCTCCCTG 945
 QY 260 lyAla-----GlnGluGluProGlyLeuArgValAlaGlyAspAsnGlyIleP 277
 Db 946 GGGCCCTTGGTGGCCACGACGCGGCTGGTGGCTCGTCCGCTGCTGGTGGCGGATCG 1005
 QY 277 roTrpSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluLeuVala 296
 Db 1006 CTTGGACCCAGGAGACAGGAGGCTCTCCAGCCCTTCTGCGACTTTCAGAAATCGTAG 1065
 QY 296 spValSerIleAsnGlnAlaProArgValGlyProAlaGlyProAlaGlyHisArgLeuValThrV 316
 Db 1066 ACATTAGCATCAAGCAGGCGCGCGCTTGGCCCGGAGAGACCGCCTGGTCACTG 1125
 QY 316 alThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuS 336
 Db 1126 TTACCAGGACAGACAAACAGATTTTAGAGCCCGAGTTCCAGGCGCTGCCAGGCTCTGT 1185
 QY 336 exPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheC 356
 Db 1186 CGTTCGTGGCGCTCGTGGAGCGGCTACTTCCGGGTGACACGAGGACTCCAGCACTTCTTCT 1245
 QY 356 yslsGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProI 376
 Db 1246 GCAAGGAGGTG---GACCCGAGGCTGCTGGAGAAAGTGGCCGAGAGTGCACGGGCCCA 1302
 QY 376 leThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrI 396
 Db 1303 TCACCTTGGACTTTGCCATCACAGCTCAAGACTGGGGCTCAGCTCTGGCTGCTATG 1362
 QY 396 leLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrP 416
 Db 1363 TTTCTCCGCCCATCCCGGAGCTTTGACAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1422
 QY 416 roLeuGlyProAspTyrLysGlyCysLeuLeuArgGlnAspProSerGlyAlaPheSerL 436
 Db 1423 CCCTTGTCTGTATTAAAGGCTGCTCTCATCGCGCAGCCCGGAGGAGGAGGAGGAGGAG 1482
 QY 436 euValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsn 456
 Db 1483 TGGTGGCTTCAGCGGACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1542
 QY 456 erGlyLeuArgValAspGlyAlaAlaLeuThrSerCysCysAlaProArgProL 476
 Db 1543 GGGGGCTGCAGCTAGATGGGTGGCGAGTGCAGCTTCTGCTGCTGCTGCTGCTGCTGCTG 1602
 QY 476 ysGluLysSerAsnLeuLeuValValArgGlyCysAsnProAlaProAlaProGlyC 496
 Db 1603 AAGAAAGTCCAACTGATTGTGTGTCAGAGAGTGCAGCCCTCAGCTTCTGCTGCTGCTGCTG 1662
 QY 496 ysSerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAsp 515
 Db 1663 TTCAGCCCCCAATCCCAATACAGCTGAGTCAATGATGATGATGATGATGATGATGATG 1722
 QY 515 erLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerA 535
 Db 1723 GCCTGGAGTGGCATGAGAACCTGGGCGATGGTCTCTTCCCAAGATTTACCGGGGCTGTC 1782
 QY 535 rGa-rGluValValAspGlyGluThrHisAspSerGluValLeuLysValMetAsp 555
 Db 1783 GCCATGAGTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1842
 QY 555 erArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnVal 575
 Db 1843 CCAAGCACAGAACTGATGAGTCTATTCCTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1902

QY 575 erTyrProHisLeuValLeuHisGlyValCysMetAlaGlyAspSerIleMetValG 595
Db 1903 CGTACCGGATCTCGTCTCCACGGCGTGTGATGGCTGGAGACAGACCATGGTGC 1962
QY 595 InGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValS 615
Db 1963 AGGAATTTGTACCTGGGGGCGCATAGATGTATCTGCGAAACGTGGCCACCTGGTGC 2022
QY 615 erAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluA 635
Db 2023 CAGCCAGCTGGAGCTGAGGTGGTCAACAGCTGGCTACGCCCTCAACTATCTGGAGG 2082
QY 635 spLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyG 655
Db 2083 ACAAGGCGCTGCCATGTGCAATGTCTCTGCCCGGAAGGTGCTCTCGTGGAGGGGG 2142
QY 655 lYAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuS 675
Db 2143 CTGATGGGAGCGCCCTCTTCATCAAGCTGAGTGACCTGGGGTTCAGCCCGCTGTGTAA 2202
QY 675 erLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaG 695
Db 2203 GCCTGGAGATGCTACCGACAGATCCCTGGGTGGCCCCCGAGTGTCTCCGGAGGGCG 2262
QY 695 InThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrTrpGluValPheG 715
Db 2263 AGACACTTAGCTTGAAGCTGACAAGTGGGGCTTCGGCGCCACGGTCTGGGAAGTGTATA 2322
QY 715 InArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluA 735
Db 2323 GTGGCGTCCACCATGCCCATCATGTCCTAGTATCCCTGCTAAGAACTCCAAATTTATGAGG 2382
QY 735 spGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysM 755
Db 2383 ACCGCGACAGCTGCGGCCCCCAAGTGGACAGAGCTGGCCCTGCTGATTCACACAGTGA 2442
QY 755 etAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyL 775
Db 2443 TGGCTATGAGCGGCTCCAGAGGCGCTCTTACGAGCGCTCATTCGTGACCTCATAGTC 2502
QY 775 euIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgA 795
Db 2503 TCATCTCTTACAGTATGAGTCTCTCTCAGACACACC---TGGTCCCTGGCACTCGTG 2559
QY 795 spGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluG 815
Db 2560 ATGGCTGTGG---AATGGTCCCGAGCTCTATGCTGCCAGACCCCGACGATCTTCGAGG 2616
QY 815 luArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuC 835
Db 2617 AGAGACACCTCAAGTACATCTCACAGCTGGGCAAGGGCTTCTTTGGCAGCGTGGAGCTGT 2676
QY 835 ysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnH 855
Db 2677 GCCCTATGACCGCTAGCGACAAATACAGTGGCTCTGGTGGCGGTGAACAGCTGCAGC 2736
QY 855 isSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLysAlaLeuH 875
Db 2737 ACAGCGGGCCAGACAGCAGGGGACTTTCAGCGGGAGATTTCAGATCTCTCAAGACAGC 2796
QY 875 isSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuA 895
Db 2797 ACAGTATTTTCATGTCAAGTATCGTGTGTGTCAGTATGGCCCGGCCCGCAGAGCCCTG 2856
QY 895 rgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg 915
Db 2857 CGCTGGTCATGGATACCTGCCAGCGCTCTTGGCGCATCTTCTGCGCGGCAACCGGG 2916
QY 915 lYLeuHisThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrL 935
Db 2917 GCCTCGATGCCAGCGGCTCTTCTTATTCTTCGACAGATCTGCAAGGGCATGGAGTACC 2976

QY 935 euGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerG 955
Db 2977 TGGGCTCCCGCCCGCTCGTGCACCGGACCTGGCCGCCGAAACATCTCGTGGAGAGCG 3036
QY 955 luAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspT 975
Db 3037 AGSCACACGTCAGATCGCTGACTTGGCCCTAGCTAAGCTGCTGCCCTTGACAAAGACT 3096
QY 975 yTrpValValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuS 995
Db 3097 ACTACGTGTGCGGAGCCAGGCGCAGCCCAATTTCTTGGTATGCCCGCAATCTCTCT 3156
QY 995 erAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluL 1015
Db 3157 CGACAAACATCTTCTCTGCCAGTCAGAGCTTGGAGCTTGGGGTCTGCTCTGTAGAGC 3216
QY 1015 euPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyP 1035
Db 3217 TCTTACCTACTGCGACAAAGCTGCGGCCCTCGGCCAGTTCCTGCGAGTATGGAT 3276
QY 1035 roGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgArgL 1055
Db 3277 GTGAGCGGATGTCCCGCCCTCTGCGGCTCTTGGAACTGCTGGAGGAGGCCAGAGGC 3336
QY 1055 euProProProProThrCysProThrGluVal 1065
Db 3337 TGCCGGCGCTCTTGTGCTGCCCTGTGAGGTG 3368

RESULT 4

US-09-003-289-1
; Sequence 1, Application US/09003289
; Patent No. 5916792
; GENERAL INFORMATION:
; APPLICANT: Civin, Curt I.
; APPLICANT: Small, Donald
; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson P.C.
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/003,289
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/357,598
; FILING DATE: 15-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/033001
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-09-003-289-1

Alignment Scores:

Db 408 TGGTT-CCCCCAGCCACATCTTCTCCGTGGAGGATGCCAGCACCCCAAGTCTCTGTGTA 466
 Qy 100 rArgLeuArgPheThrProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLe 120
 Db 467 CAGGATTCTGCTTTACTTCCCAATTTGGTTGGCTGGAGAGTCCACCGCTTCGGGCT 526
 Qy 120 uArglyAspLeuThrSerAlaIleuAspLeuHisValLeuGluHisLeuPheAlaG 140
 Db 527 AGCAGAGGATTTGGCAGGCTGATCTTACCTGCCAGTCTCGAGACCTCTTTGCCCA 586
 Qy 140 nHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGln 160
 Db 587 GCACCGCAGTACCTGGTGGTGGGCGCTCCCGCTGGCTCAGTCTCAGAGCAGGG 646
 Qy 160 yGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGln 180
 Db 647 TGAGTGTCTCAGCCTGGCGTGTGGACCTGGCGCGGATGGCGGAGAGCAGGCCAGCG 706
 Qy 180 gProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAs 200
 Db 707 GCGGGAGAGCTGCTGAAGACTGTGAGTCAAGGCTGCCTACCCCAAGCTCGCGGA 766
 Qy 200 pValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAl 220
 Db 767 CTGTATCCAGGGCTGAGCTTCGTGACGGGGGCGGTATTTCGGAGGACGGTGGAGAGCC 826
 Qy 220 aLeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleL 240
 Db 827 CTGTGCG-CCGGTGGCGCTGCTGACGACGACGCGCACTCGCTCATGGCCAAAGTATCA 885
 Qy 240 euAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260
 Db 886 TGACCTGGAGCGCTGGATCCAGCGGGCGCGCGAGACTTCCACGTGGGCTCCCTG 945
 Qy 260 lYAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleP 277
 Db 946 GGGCCCTTGGTGGCCACGACGGGTGGGCTCGTCCGCTGGTGTGACGGCGGCATCG 1005
 Qy 277 roTrpSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleVal 296
 Db 1006 CTTGGACCCAGGAGAACAGAGGTCTCCAGCCCTTCTCGACTTTTCCAGAAATCGTAG 1065
 Qy 296 spValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrV 316
 Db 1066 ACATTAGCATCAACAGCGCCCGCGCTGGCCCGCGGAGACACCGCTGTCTACTG 1125
 Qy 316 alThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuS 336
 Db 1126 TTACCCAGGACAGACAACAGATTTTAGAGGCGGAGTTCCACAGGCTGCGCGAGGCTCTGT 1185
 Qy 336 erPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheC 356
 Db 1186 CGTTCGTGGCGCTCGTGACCGCTACTTCCGCTGACCCAGCAGCTCCAGCACTTCTTCT 1245
 Qy 356 yLysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProI 376
 Db 1246 GCAAGGAGGTG---GACCCGAGGCTGCTGGAGGAGTGGCCGAGAGTCCACGCGCCCA 1302
 Qy 376 leThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrI 396
 Db 1303 TCACCTGGACTTTGGCCATCAACAGCTCAAGACTGGGGGCTCAGTCTCGCTCTCTATG 1362
 Qy 396 leLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrP 416
 Db 1363 TTCTCCGCGCATCCCGCAGGACTTTCAGCTTCTCCTCCTCAGTCTGTGTCCAGAAC 1422
 Qy 416 roLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerL 436
 Db 1423 CCCTTGGTCTGATTATTAAGGCTGCTCCATCCGCGCAGCCCAACAGAACTTCTCTTC 1482
 Qy 436 euValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaCysTrpAsnS 456
 Db 1483 TGGTTGGCTCAGCGCACCCACAGAGTCTTTCGAGAGCTCTCTGGCAACCTGCTGGGATG 1542

Qy 456 erGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProL 476
 Db 1543 GGGGCTGCACGTAGATGGGTGGCAGTGAACCTCACTTCTCTGTATATCCACAGACCA 1602
 Qy 476 yGlyLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyC 496
 Db 1603 AAGAAAGTCCAACTGATTGGTGGTCCAGAGAGGTGACAGCCCAACCATCATCTTGG 1662
 Qy 496 ySerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspS 515
 Db 1663 TTCAGCCCAATCCCAATACAGCTGAGTCAGATGACATTTTCAAGATCCCTGTGTACA 1722
 Qy 515 erLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerA 535
 Db 1723 GCTGTGAGTGGCATGAGAACTGGGCGCATGGTCTTCAAGATTTTACCGGGGCTGTG 1782
 Qy 535 rArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspS 555
 Db 1783 GCATGAGTGGTGGATGGGAGGCGCCGAAAGACAGAGGTGCTGTGAAGGTTCATGATG 1842
 Qy 555 erArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValS 575
 Db 1843 CCAAGCACAAGAACTGCATGGAGTCACTTCTGGAAGCAGCAGCTTGTATGAGCAAGTGT 1902
 Qy 575 erTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValG 595
 Db 1903 CGTACCGCATCTCGTGTCTCCAGCGGTGTGTCATGGTGGAGACAGCAGCATGTGTG 1962
 Qy 595 lGlnPheValTyrLeuGlyAlaIleAsnMetTyrLeuArgLysArgGlyHisLeuValS 615
 Db 1963 AGGAATTTGACACTGGGGGCGCATGACATGATATCTGCCAAACAGTGGCCACTGTGTG 2022
 Qy 615 erAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluA 635
 Db 2023 CACCCAGCTGGAAGCTGCAGTGGTCAACAGCTGGCTTACGCCCTCAACTATCTGAGG 2082
 Qy 635 spLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyG 655
 Db 2083 ACAAGGCTGTCCCATGGCAATGTCTCTGCCGGAAGGTGCTCTCGCTGGGAGGGGG 2142
 Qy 655 lyAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuS 675
 Db 2143 CTGATGGAGCCCGCCCTTTCATCAAGCTGAGTGAACCTGGGTGAGCCCGCTGTGTATA 2202
 Qy 675 erLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaG 695
 Db 2203 GCCTGGAGATGCTCACCGACAGGATCCCTGGGTGGCCCGAGTGTCTCGGAGGCGC 2262
 Qy 695 lThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheG 715
 Db 2263 AGACACTTAGCTTGAAGCTGCAAGTGGGCTTTCGGCGCCACCGGTCTGGGAAGTGTATA 2322
 Qy 715 lNArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluA 735
 Db 2323 GTGGCTGCACATGCCCATCAGTCCCTAGATCTCTGTAGAAACTCAATTTTATGAGG 2382
 Qy 735 spGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysM 755
 Db 2383 ACCGGCAGCAGCTGCGGCCCCCAAGTGGACAGAGTGGCCCTGTCTGATTCACAGTGCA 2442
 Qy 755 etAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyL 775
 Db 2443 TGGCTATGAGCCGCTCCAGAGGCGCTCTTACGAGCGCTCATTTCTGACCTCAATATGC 2502
 Qy 775 eulleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgA 795
 Db 2503 TCATCTCTTCAGACTATGAGCTCTCTCAGACCCACAC---TGGTGGCTTGGCACTCGTG 2559
 Qy 795 spGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluG 815
 Db 2560 ATGGCTGTGG---AATGGTGGCCAGCTTATGCTGCGCAAGACCCCAAGTCTTCGAGG 2616

QY 815 luArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuC 835
Db 2617 AGAGACACCTCAAGTATCATCTCAGCTGGGCAAGGCTTCTTTGGCAGCGTGAGCTGT 2676
QY 835 YeArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuH 855
Db 2677 GCGCTATGACCGCTAGGCGCAATACAGTGCCCTGGTGGCGGTGAACACAGCTGCAGC 2736
QY 855 iSerValProAspGlnGlnArgAspPheGlnArgGluLeuGlnLeuLysAlaLeuH 875
Db 2737 ACAGCGGGCAGACACAGCAGGAGCTTTTCAGCGGGAGATTCAGATCCTCAAGACAGC 2796
QY 875 iSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuA 895
Db 2797 ACAGTGATTTCAATGTCAAGTATCGTGGTGCAGCTATGCGCGGCGCGCCAGCGCTG 2856
QY 895 rglLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg 915
Db 2857 CGTGGTCAATGAGTACCTGCCAGCGGCTGCTTGCAGCTTCCGCGAGCGCACCGG 2916
QY 915 lLeuHisThrAspArgLeuLeuLeuPheAlaTyrGlnIleCysLysGlyMetGluTyrL 935
Db 2917 GCCTCGATGCCAGCGCTCTCTCTATTCCTCGCAGATCTGCAAGGCGATGGATACC 2976
QY 935 euGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerG 955
Db 2977 TGGGTCCCGCGCTGCGTGCACCGCGACTGCGCGCGCGAAACATCTCTGTTGGAGAGCG 3036
QY 955 luAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuProLeuGlyLysAspT 975
Db 3037 AGGCACAGTCAGATCGCTGCTTCCGCTAGCTAAGCTGCTGCGGCTTGCACAAAGACT 3096
QY 975 YTYrValValArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuS 995
Db 3097 ACTAGTGTGTCGCGAGCGAGCGCCAGCCCATTTCTGATGATGCCCGCGATCTCTCT 3156
QY 995 erAspAsnIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluL 1015
Db 3157 CGGACAACTCTCTCGCCAGTCAGAGCTGCGAGCTTCCGGTCTGCTGTACGAGC 3216
QY 1015 euPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyP 1035
Db 3217 TCTTCACTACTGCGACAAAGCTCGACCCCTCGCGCGCTTCCGCGATGTATGGAT 3276
QY 1035 roGluArgGluGlyProProLeuCysArgLeuLeuGluLeuLeuAlaGluGlyArgArgL 1055
Db 3277 GTGAGCGGGATGTCGCCCGCTCTGCCGCTCTTGGAACTGCTGAGGAGGCGCCAGAGC 3336
QY 1055 euProProProThrCysProThrGluVal 1065
Db 3337 TGCGGCGCGCTCTGTGCTGCGCTGTGAGGTG 3368

RESULT 6

US-09-046-158A-21
; Sequence 21, Application US/09046158A
; Patent No. 6187552
; GENERAL INFORMATION:
; APPLICANT: Robert, Steven L.
; APPLICANT: Kaytes, Paul S.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING INHIBITORS OF
; TITLE OF INVENTION: JAK2/CYTOKINE RECEPTOR BINDING
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pharmacia & Upjohn Co., Intellectual Property
; STREET: 301 Henrietta Street
; CITY: Kalamazoo
; STATE: MI
; COUNTRY: USA
; ZIP: 49001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/046,158A

FILING DATE:
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Darnley Jr., James D.

REGISTRATION NUMBER: 33,673

TELEPHONE: 616/833-2210

TELEFAX: 616/833-8897

TELEX: 224401

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 3435 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-09-046-158A-21

Alignment Scores:

Pred. No.: 1.2e-234 Length: 3435
Score: 2629.00 Matches: 534
Percent Similarity: 66.21% Conservative: 193
Best Local Similarity: 48.63% Mismatches: 340
Query Match: 44.86% Indels: 31
DB: 3 Gaps: 14

US-09-397-967A-16 (1-1099) x US-09-046-158A-21 (1-3435)

QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProG1 39
Db 111 TCCAGTTCTTCAGT-----GTATCTTTACCATTCCTTGGAAATCTGAGGCAGA 161
QY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaLay 59
Db 162 TTATCTGACCTTTCATCTGGGAGTATGTTCGAGAGAAATCTGATTGCTGCTTCTAA 221
QY 59 sAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSe 79
Db 222 AGCTTGATGATACACCTGTGATCATATATGTTTCTTTAATGATGAACAGAAAG 281
QY 79 rCysTrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa 99
Db 282 GATCTGTATCCACCCACCATGCTTCCATATAGATGAGTCAACAGGCATATGTA 341
QY 99 lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr 117
Db 342 CTACAGAAATAAGATTTTACTTTCTCTGTTGTTTATGAGTGGCAGCAACAGAGCTATCG 401
QY 117 gPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137
Db 402 GCATGGAATATCTCGAGGTGCTGAGCTCTCTTCTTGTATGACTTTGTCATGCTTACCT 461
QY 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157
Db 462 CTTTGTCTCAGTGGCGCATGATTTTGTGTCATGATGATGATGATGATGATGATGATGATG 521
QY 157 sGluGlnGlyPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl 177
Db 522 AACACAGGAAGAATGCTCTCGGATGCGAGTGTGTAGATATGATGATGATGATGATGATGATG 581
QY 177 nAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSe 197
Db 582 CGATCAACCCCTGCGCATCTATTAACCTATCAGCTACAAGACATCTTACCAAAATG 641
QY 197 rLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgIleArgThrVa 217
Db 642 TATTCGAGCAAGATCCAGACTATCATATTTTGGACAGGAAGCAATAGGTACAGATT 701
QY 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlay 237

b	702	TGCAGATTATTCAGCAAACTTCCAGCAATGCGAAGGCCACTGCCAGAACTTTGAACCTTAA	761
y	237	sTyrlleLeuAspLeuGluArgLeuHieProAlaAlaThrThrGluThrPheArgValGI	257
b	762	GTAATCTATAAATCTGGAAACTCTGCAGTCTGCCCTTCTACACAGAGAAATTTGAAGTAAA	821
y	257	yLeuProGlyAla-----GlnGluGluProGlyLeuArgValAlaGlyAs	273
b	822	AGAACCCTGGAAGTGTCCTTCAGGTGAGGAGATTTTTGCAACCATATATAAACAATGGA	881
y	273	pAsnGlyIleProTrpSer-----AsnAspGluLe	284
b	882	CGGTGGAATTCAGTGTCAAGAGGGAACATAAAGAAAGTGAGACACTGCAGAACACGA	941
y	284	uPheGlnThrPheCysAspPheProGluLeuValAspValSerIleAsnGlnAlaProAr	304
b	942	TTTACAGTTATATTCGATTTTCTTAATATATTTATGATGTCAGTATTAAGCAACGCA	1001
y	304	gValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLe	324
b	1002	AGAGGGT---TCAAATGAAAGCCGAGTTGTAATCTCCATGAAGAGATGGTAAANAATCT	1058
y	324	uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTY	344
b	1059	GGAAATTTGAACCTTAGCTCATTAGGAGAGCTTTGTCTTCGTGTCATTAAATGATGATA	1118
y	344	rPheArgLeuIleCysAspSerArgHisTyPheCysLysGluValAlaProProArgLe	364
b	1119	TTATAGATTAACTGCAGATGCACATCATTAACCTCTCTAAAGAAAGTAGCACCTCCAG	1178
y	364	uLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLy	384
b	1179	GCITGAAAAATATACAAAGCAACTGTCTAGTCCCAATTTGATGGATTTTGCCTATTAGTAA	1238
y	384	sLeuIySaLaAlaGlySerLeuProGlyThrTyrlleLeuArgArgSerProGlnAspTY	404
b	1239	ACTGAAGAAACGAGTAAATCMAGTCGACGTATGATCTTCGATCGAGCTCTAAGGACTT	1298
y	404	rAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrylsGlyCy	424
b	1299	TAATAAATATTTTGTGCTTTGCTGTCGAGCGAGAAATGTCAATTAATAAACAACGT	1358
y	424	sLeulleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisar	444
b	1359	TTTGATTAACAAAAATGAGATGAAGAGTCAACCTCAGTGGGACAAAGAAAGAACTTCAG	1418
y	444	sSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAl	464
b	1419	CAGTCTTAAGACATCTTTTGAATTTGTTACCAGATGGAACCTGTCGCTCAGACAATATA	1478
y	464	aLeuTyrlleThrSerCysAlaProArgProGlyGluLysSerAsnLeulleValVa	484
b	1479	TTTCCAGTTTACTAAATGCTGTCCCCCAAGCCAAAAGATAAATCAAACTCTAGTCTT	1538
y	484	Iarg--ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAlaLe	503
b	1539	CAGAACGAATGCTGTTCTGTATGCCAACCTCCAAACATTAACAGAGGCTACTCNAT	1598
y	503	uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGl	523
b	1599	GAACCAAAATGGTGTTCACMAAATCAGAAATGAAGATTTGATATTAATGAAGCCTTG	1658
y	523	yHisGlySerPheThrLyIlePheArgGlySerArgArgGluValValasp---GlyGI	542
b	1659	CCAAGGCACCTTTTACAAAGATTTTAAAGGCGTACGAAGAGAGTAGGAGACTACGGTCA	1718
y	542	uThrHisAspSerGluValLeuLeuIlysValMetAspSerArgHisArgAsnCysMetGI	562
b	1719	ACTGCATGAACAGAAAGTTCTTTTAAAGTTCTGTGATAAGCACACAGAAACTATTAGA	1778
y	562	nSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrlleHisLeuValLeuLe	582

QY	284	uPheGlnThrPheCysAspPheProGluLeuValAspValSerIleAsnGlnAlaProAr	304	Db	2355	GAATGTATGTGCAAAATAATTTCTCTTATCAGAGAAGACAGACAGGAATCC	2414
Db	1278	TTTACAGATTATATTCGATTTTCCATAATATTTATGATGATGATTAAGCAACCA	1337	QY	659	oPheIleIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGlyMetLe	679
QY	304	gValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLe	324	Db	2415	TCCTTTTCATCAAACTTAGTGATCTCGCATTAAGTATACAGTTTGGCAAGACATCT	2474
Db	1338	AGAGGGT--TCAAATGAAGCGGAGTGTACTATCCATTAAGCAAGATGGTAAAAATCT	1394	QY	679	uThrAspArgIleProTTPValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLe	699
QY	324	uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTy	344	Db	2475	TCAGAGAGAAATACCATGGGTACCACTGAATGCATTTGAAAAATCTCTAAATTTTAA	2534
Db	1395	GGAAATTTGAATTTAGCTATTAAAGGAAGCTTTGCTTTTCGTGTCATTAATTGATGATA	1454	QY	699	uGluAlaAspLysTyrGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAl	719
QY	344	rPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProArgLe	364	Db	2535	GGCAACAGACAAATCGAGTTTGGTACCACTTTGGGAAATCTCGAGTGGAGGAGATAA	2594
Db	1455	TTATAGATTAACTGCAGATGCACATCATCTCTGTAAGAGTAGCACCTCCAGCCGT	1514	QY	719	ahIstIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLe	739
QY	364	uLeuGluGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLy	384	Db	2595	ACCTCTAAGTGTCTGTGATCTCTCAAGAAAGCTACAAATTTTATGAAGATAGCATCAGCT	2654
Db	1515	GCTTGAAATATACAAAGCAACTGTCTATGGCCCAATTTTCGATGATTTTGCCATTAGTAA	1574	QY	739	uProAlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPr	759
QY	384	sLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTy	404	Db	2655	TCCTGCACCAAGTGGCGAGAATTTAGCAACCTTTATAATAATTTGTATGGATTATGAACC	2714
Db	1575	ACTGAAGAAAGCAGGTAATCAGACTGCTGTATGATCTCGATGAGTCTTAAGACTT	1634	QY	759	oGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs	779
QY	404	rAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCy	424	Db	2715	AGATTTTCAGGCCCTTCTTTCAGAGCCATCATACAGATCTTAACAGTTTGTTTACTCCAGA	2774
Db	1635	TAATAAATATTTTTCATCTTCTGTCGAGCGAGAAATGTCTGTAATATTAACACTG	1694	QY	779	pTyrGluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCy	798
QY	424	sLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAr	444	Db	2775	TTATGAACCTATTAAACAGAA---AATGACATGTTTACCAATATGAGGATAGTCCCTTGGG	2831
Db	1695	TTTGATTACAAAATGAGATGAGAGTACACCTCAGTGGGACAAAGAACTTCAG	1754	QY	798	sValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLe	818
QY	444	gSerLeuArgGluLeuLeuAlaLaCysTrpAsnSerGlyLeuArgValAspGlyAlaAl	464	Db	2832	GTITTTCTGGTGCC---TTTGAAGACCGGATCTCTACACAGTTTGAAGAGAGACATTT	2885
Db	1755	CAGTCTTAAAGATCTTTTGAATTTTACCAGATGGAACCTGTTTCGCTCAGACAAATATAAT	1814	QY	818	uLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAs	838
QY	464	aLeuTyrLeuThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIleValva	484	Db	2886	GAATTTTCTACAGCAACTTGGCAAGGGTAATTTTGGAGGTGGAGATGTCGCGGTATGA	2945
Db	1815	TTTCCAGTTTACTAAATGCTGTCCCCAAAGCCAAAGATAAATCAAACTTCTAGTCTT	1874	QY	838	pProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPr	858
QY	484	laIq---ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysCysAlaLe	503	Db	2946	CCCTCTACAGCAACACTCTGGGAGGTGTCGCTGTAAATAAGCTTCAGCATAGTACTGA	3005
Db	1875	CAGAACAAATGGTGTCTGATGTACCACTCCACCAATTCAGAGGCGCTACTCATAT	1934	QY	858	oAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPh	878
QY	503	uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGl	523	Db	3006	AGACCACTTAAGAGACTTTTGAAGGGAAATTTGAAATCTCTGAAATCTCCATACAGATGACAA	3065
Db	1935	GAACCAAAATGGTGTTCACAAAATCAGAAATGAAGATTTGATATTTAATGAAGCCCTTGG	1994	QY	878	eIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMe	898
QY	523	yHisGlySerPheThrLysIlePheArgGlySerArgArgGluValValAsp---GlyGl	542	Db	3066	CATTGTAAGTACAAAGGAGGTGCTACAGTGTGGTGGCGGTAAATCTTAAATTAATATAT	3125
Db	1995	CCAAAGGCACCTTTTACAAAGATTTTAAAGGCGTACGAGAGAGATGAGAGACTACGGTCA	2054	QY	898	tGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHi	917
QY	542	uThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGl	562	Db	3126	GGAATATTTTACCATTATGGAAGTTTACGAGACTATCTTCAAAAACATAAAGAACGATAGA	3185
Db	2055	ACTGCATCAACACAGAGTCTTTTAAAGTTCTGATAAAGCACACAGAGAACTATTACAGA	2114	QY	917	sThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAl	937
QY	562	uSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLe	582	Db	3186	TCACATAAACTTCTGCAGTACACATCTCAGATATGCAAGGTATGAGTATCTTTGGTAC	3245
Db	2115	GTCTTTCTTGAAGCAGCAAGTATGATGAGCAAGCTTCTCACAAGCAATTTGGTTTAAA	2174	QY	937	aArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHi	957
QY	582	uHisGlyValCysMetAlaGlyAsp---SerIleMetValGlnGluPheValTyrLeuGl	601	Db	3246	AAAAAGGTATATCCACAGGGATCTGGCAACGAGAAATATATTTGGTGGACGACGACAG	3305
Db	2175	TTATGGAGTATGTCTGTGGAGCAGCAATATTTCTGTTTCTGAGGATTTGTAAATTTGG	2234	QY	957	sValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrVa	977
QY	601	yAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLysLeuGl	621	Db	3306	AGTTTAAATTTGGTATTTTGGTTTAAACCAAGTCTGCCACAGACAAAGAATACTATAA	3365
Db	2235	ATCAGTACATATCTGAAAGAGATAAATAATTTGTTATATATATTTATGGAACCTTGA	2294	QY	977	lValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluLeuSerLeuSerAspAs	997
QY	621	nValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGl	641	Db	3366	AGTAAAGAACTCTGTGTAAGTCCCATATTTCTGTGTATCTCCAGATCTCAGACAGAG	3425
Db	2295	AGTTGCTAAACAGTTGGCATGGCCCATTTTCTGAAGAGAAACACCCCTTATTCATGG	2354	QY	997	nIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPhePh	1017
QY	641	yAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGly-----AspGlyAsnPr	659				

Db 3426 CAAAGTTTCTGGGCTCAGATGTTGGAGCTTTGGAGTGGTCTGTATGAATTTTCAC 3485
QY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluAr 1037
Db 3486 ATACATTGAGAGAGTAAAGTCCACCAGGGNAATTTATGCGTATGATTGGCAATGACAA 3545
QY 1037 gGluGlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuPr 1056
Db 3546 ACAAGCAGATGATCGTGTCTCCATTGATAGAACTTTTGAAGAATAATGGAAGATTACC 3605
QY 1056 oProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGl 1076
Db 3606 AAGACACGATGATGCCAGATGAGATCATATGATCATGACAGATGCTGGAACAATAA 3665
QY 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093
Db 3666 TGTAATCAACGCCCTCTCTTAGGATCTAGCTCTTCAGTGGATCAANAATA 3717

RESULT 8

US-09-196-480-1
; Sequence 1, Application US/09196480
; Patent No. 6019966
; GENERAL INFORMATION:
; APPLICANT: Coleman, Roger
; APPLICANT: Stuart, Susan G.
; TITLE OF INVENTION: A NOVEL HUMAN JAK2 KINASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,480
; FILING DATE:

CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/567,508
; FILING DATE: 05-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0049US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4482 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY: Placenta
; CLONE: 179527
US-09-196-480-1

Alignment Scores:
Pred. No.: 1.03e-233 Length: 4482
Score: 2621.00 Matches: 532
Percent Similarity: 66.03% Conservative: 193
Best Local Similarity: 48.45% Mismatches: 342
Query Match: 44.73% Indels: 31
DB: 3 Gaps: 14

US-09-397-967A-16 (1-1099) x US-09-196-480-1 (1-4482)
QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProGl 39
Db 447 TCCAGTTCTTCAGGT-----GTATCTTTACCATTCCTTGGGAAATCTGAGGCAGA 497
QY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluLeuLeuLeuValArgAlaAla 59
Db 498 TTATCTGACCTTTCCATCTGGGAGTATGTTGGAGAAGAAATCTGTATTGCTCTCTAA 557
QY 59 sLaCysGlyLeuLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPhe 79
Db 558 AGCTTGTGTATCACACCTGTGTATCATATAATGTTTCTTTAATGAGTGAACAGAAAG 617
QY 79 rCysTrpPheProProSerHisIlePheCysIleGluAspValAlaAspThrGlnVal 99
Db 618 GATCTGGTATCCACCAACCATGTCTTCATATAGATGAGTCAACAGGCATATGTACT 677
QY 99 lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHis 117
Db 678 CTACAGATATAGATTTTACTTCTCTGTTGGTATTCAGTGGCAGCAACAGAGCCTACG 737
QY 117 gPheGlyLeuArgLysAspLeuThrSerAlaLeuLeuAspLeuHisValLeuGluHis 137
Db 738 GCATGGAATATCTCGAGGTGCTGAAGCTCCTCTTCTTGATGACTTTGTCTTACT 797
QY 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMet 157
Db 798 CTTTGTCTCAGTGGCGCATGATTTTGTGATGGATGGATAAAAGTACCTGTGACTCATGA 857
QY 157 sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlu 177
Db 858 AACACAGGAAGAATGCTCTTGGGATGACAGTGTATGATATGATGATAGCAAGAGAAA 917
QY 177 nAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProSe 197
Db 918 CGATCAAAACCCACTGGCCATCTATACTCTATCAGCTACAGACATCTTACCACAATG 977
QY 197 rLeuArgAspValIleGlnGlnAsnPheValThrArgArgIleArgThrVa 217
Db 978 TATTTCAGCAAAAGATCCAAGACTATCATATTTTTCACAGGAAGCAATTAAGTACAGATT 1037
QY 217 lValLeuAlaLeuLeuProCysGlyArgGluLeuProGlyArgProTyrAlaLeuMetAla 237
Db 1038 TCGCAGATTATTTCAGCAATTCAGCAATGCAAAAGCCACTCCCAAGAACTTGAACCTTAA 1097
QY 237 sTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal 257
Db 1098 GTATCTTATAAATCTGGAACTCTGCAGTCTGCCTTCTACACAGAAATTTGAAGTAAA 1157
QY 257 yLeuProGlyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAs 273
Db 1158 AGAACCTTGAAGTGGTCTCTTCAGGTGAGGAGATTTTTCACACCATATATAATACTGAAA 1217
QY 273 pAsnGlyIleProTrpSerSer-----AsnAspGluLe 284
Db 1218 CGTGGAAATTCAGTGGTCAAGAGGAAACATAAGAAAGTGCAGACACTGACAGAACAGGA 1277
QY 284 uPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProAr 304
Db 1278 TTTACAGTTATATGCGATTTTCCATAATATATGATGTCAGTATTAAGCAAGCAACCA 1337
QY 304 gValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIle 324
Db 1338 AGAGGGT---TCAAATGAAAGCCAGTTGTAACCTATCCATAAGCAAGATGTTAAATCT 1394
QY 324 uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTy 344
Db 1395 GGAAATGGAATCTAGCTCATTAAGGGAAGCTTTGTCTTCGTCAATTAATGATGATA 1454
QY 344 rPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProArgLe 364
Db 1455 TTATAGATTAACTGCAGATGCACATCATTTACCTCTGTAAAGAAAGTAGCACCTCCAGCCGT 1514


```
Db 3666 TGTAAATCAACGCCCTCTCTTAGGATCTAGTCTTCGAGTGATCAAAATA 3717
RESULT 9
US-09-972-800A-15
; Sequence 15, Application US/09972800A
; Patent No. 6534277
; GENERAL INFORMATION:
; APPLICANT: Hancock, W.
; APPLICANT: Ozkavnak, B.
; TITLE OF INVENTION: ROLES OF JAK/STAT FAMILY MEMBERS IN TOLERANCE
; TITLE OF INVENTION: INDUCTION
; FILE REFERENCE: 7853-192
; CURRENT APPLICATION NUMBER: US/09/972,800A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US/09/549,654
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 5117
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (478)...(3876)
US-09-972-800A-15

Alignment Scores:
Pred. No.: 1,27e-233 Length: 5117
Score: 2621.00 Matches: 533
Percent Similarity: 66.12% Conservative: 193
Best Local Similarity: 48.54% Mismatches: 341
Query Match: 44.73% Indels: 31
DB: 4 Gaps: 14

US-09-397-967a-16 (1-1099) x US-09-972-800A-15 (1-5117)

Qy 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProGly 39
Db 585 TCCAGTCTTCAGGT-----GTATCTTTACCATTTCCCTTGGGAATCTGAGGCAGA 635
Qy 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLy 59
Db 636 TTATCTGACCTTCCATCTGGGAGTATGTTGCAGAGAAATCTGATGCTGCTTCAA 695
Qy 59 sAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSe 79
Db 696 AGCTTGTGTATCACCCTGTATCATTAATATGTTTGTCTTAATGAGTGAACAGAAAG 755
Qy 79 rCysTrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa 99
Db 756 GATCTGGTATCCACCCACCATGCTTCCATATAGATGAGTCAACCGGCATATGTA 815
Qy 99 lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr 117
Db 816 CTACAGATAAGATTTTACTTCTCGTGTGATTTGCAGTGGCAGCAACAGGCCTATCG 875
Qy 117 gPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137
Db 876 GCATGGAATATCTCGAGGTGCTGAAGCTCCTCTTCTTGATGACTTTTGTCTACCT 935
Qy 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157
Db 936 CTTTGTCTAGTGGCGGATGATTTTGTGTCAGGATGATATAAAGTACTGTGATCATGA 995
Qy 157 sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl 177
Db 996 AACACAGGAAGAAATGCTCTGGAGTGGCAGTGTAGATATGATGAGATAGCCAAAGAAA 1055
Qy 177 nAlaGlnArgProGlyGluLeuLeuYsThrValSerTyrIysAlaCysLeuProProSe 197
Db 1056 CGATCAAAACCCCACTGCGCCATCTATACTCTATCAGCTACAGACATTTCTTACCAAAATG 1115
```

```
Qy 197 rLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrVa 217
Db 1116 TATTCGAGCAAGATCCCAAGACTATCATATTTTGACAGGAAGCAATTAAGGTACAGATT 1175
Qy 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLy 237
Db 1176 TCGCAGATTTATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCA 1235
Qy 237 sTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGl 257
Db 1236 GTATCTTATAATCTGGAACTCTGCAGTCTGCCTTCTACACAGAGAAATTTGAAGTAA 1295
Qy 257 yLeuProGlyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAs 273
Db 1296 AGAACCTCGAAGTGTCTCTCAGGTGAGGAGATTTTTCACCACTATAATACTGGAAA 1355
Qy 273 pAsnGlyIleProTrpSerSer-----AsnAspGluLe 284
Db 1356 CGGTGGAATTCAGTGTCAAGAGGGAACATAAAGAAAGTGAGACACTCACAGAACGGA 1415
Qy 284 uPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProAr 304
Db 1416 TTTACAGTTATATTCGATTTTCTTAATATTATTGATGTCAGTATTAGCAAGCAACCA 1475
Qy 304 gValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLe 324
Db 1476 AGAGGT---TCAAAATGAAAGCCGAGTTGTAACTATCATTAAGCAAGATGTTAAATCT 1532
Qy 324 uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTy 344
Db 1533 GGAAATTTGAATTCAGTCTTAAAGGAAGCTTTGCTTTCGTGTCATTAAATTTGATGGATA 1592
Qy 344 rPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLe 364
Db 1593 TTATAGATTAACTGCAGATGCACATCATTTACTCTCTGTAAGAGAGTAGACCTCCAGCCGT 1652
Qy 364 uLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLy 384
Db 1653 GCTTGAATAATATACAAAGCAACTGTCATGGCCCAATTTTCGATGATTTTGCATTAGTAA 1712
Qy 384 sLeuLyAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTy 404
Db 1713 ACTGAGAAAGCAGTAAATCAGACTGGACTGTATGATCTCTCGATGCGAGTCTCTAGGACTT 1772
Qy 404 rAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCy 424
Db 1773 TAATAAATATTTTGTACTTTTGTCTGTCGAGCGAGAAATGCTCATTTGAATATAACACTG 1832
Qy 424 sLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAr 444
Db 1833 TTTGATTACAAAATAATGAGATGAAGAGTACAACTCAGTGGGCAAGAAAGAACTTCAG 1892
Qy 444 gSerLeuArgGluLeuLeuAlaCysTyrPheAsnSerGlyLeuArgValAspGlyAlaAl 464
Db 1893 CAGTCTTTAAGATCTTTTGAATGTTTACCAGATGGAAACTGTTCTCGTCAGACAAATATA 1952
Qy 464 aLeuTyrLeuThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIleValVa 484
Db 1953 TTTCCAGTTTACTAAATGCTGTCCCCCAAGGCCAAAGATAAATCAAACTTCTAGTCTT 2012
Qy 484 lArg---ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAlaLe 503
Db 2013 CAGAACGAATGGTGTCTCTGATGTACCACTCACCACCACTTACAGAGGCTCTACTCAT 2072
Qy 503 uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGl 523
Db 2073 GAACCAATATGGTGTTCACAAAATCAGAAATGAAGATTTTGATATTTAATGAAGCCTTG 2132
Qy 523 yHisGlySerPheThrLysIlePheArgGlySerArgArgGluValValAsp---GlyGl 542
Db 2133 CCAAGGCACCTTTTCAAAAGATTTTAAAGCGCTACGAAGAGAGAGTAGGAGACTACGCTCA 2192
```


TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3629 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 94...3480
 US-08-097-997A-8

Alignment Scores:

Pred. No.: 1,01e-233 Length: 3629
 Score: 2619.50 Matches: 537
 Percent Similarity: 65.50% Conservative: 190
 Best Local Similarity: 48.38% Mismatches: 350
 Query Match: 33 Indels: 33
 DB: 1 Gaps: 15

US-09-397-967A-16 (1-1099) x US-08-097-997A-8 (1-3629)

QY 4 ProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSerGluAla 23
 DB 145 CCTGTACATCAAGTGGTGGATATCTCGAAGTGCTAAATTTCTGAAGCAGATAGGCCA 204
 QY 24 GlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArgLeuSerPhe 43
 DB 205 GTCTTCAAGTGATCTGTACCATTTCTTGGCAAGCTGAAGGAGATCTGAAGTTT 264
 QY 44 SerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAlaCysGlyIle 63
 DB 265 CCAAGTGGAGATATGTTGCAGAGAAATTTGTGGCTGCTCTTAAGCTTTGGTATT 324
 QY 64 LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPhePro 83
 DB 325 ACGCTGTGTATCATATAATATGTTGCGTTAATGATGAACCGAAGGATCTGTGATCCCA 384
 QY 84 ProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg 103
 DB 385 CCCAATCATGTCTCCATATAGACAGTCAACCGGCGATGATCATCTACAGGATAAGG 444
 QY 104 PheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisArgPheGlyLeuArg 121
 DB 445 TTCTACTTCCCTCATTTGGTACTGTAGTGGCAGCAGCAGCAACCTACAGATACGGAGTGCC 504
 QY 122 LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141
 DB 505 CGTGGGCTGAAGCTCCTCTCTGTATGATGACATTGTTCATGCTTACCTTTTGTCTCAGTGG 564
 QY 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGlu 161
 DB 565 CGGCATGATTTTGTTCAGGATGGATTAAGTACTGTGACCTCATGAACCTCAGGAAGAG 624
 QY 162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaArgPro 181
 DB 625 TGTCTTGGATGGCGGTGTAGACATGATGATGATGATGATGATGATGATGATGATGATGAT 684
 QY 182 GlyGluLeuLeuLysThrValSerTyrIlyAlaCysLeuProProSerLeuArgAspVal 201
 DB 685 CTGGCTGTCTATATCTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 744
 QY 202 IleGlnGlyGlnAsnPheValThrArgArgArgIleArgThrValValLeuAlaLeu 221
 DB 745 ATCCAGACTATACATTTTACCGGAGGAGGATCAGGTACAGATTTCGCAGATTTCATT 804
 QY 222 LeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAsp 241
 DB 805 CAGCAATTCAGTCAATGTAAAGCCACTGCCAGGAACCTTAACTTAACTTAACTTAACTTAA 864
 QY 242 LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal----- 256
 DB 865 CTGGAAACCCCTGCGTCTCTACACAGAACAGTTTGAAGTAAAGAAAGTCTGCAAGA 924

QY 257 GlyLeuProGlyAlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle 276
 DB 925 GGT---CCTTCAGGTGAGGAGATTTTTCACCATTAATAAATTAAGTGAACGTTGAAT 981
 QY 277 ProTrpSerSer-----AsnAspGluLeuPheGlnThr 287
 DB 982 CAGTGTCAAGAGGAAACATAGAAAGTGAAGACACTGACAGAACAGGACGATCAGTTA 1041
 QY 288 PheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyPro 307
 DB 1042 TATTGTGATTTCCCTGATATTATTGATGATTAAGCAAGCAACAG---GAATGC 1098
 QY 308 AlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGlu 327
 DB 1099 TCAATGAAGTAGAATTTGTAATCTCCATATAAAGATGTTAAAGTTTGGAGATAGAA 1158
 QY 328 PheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347
 DB 1159 CTAGCTCATTAAGAAGACCTTGTCTGTCATTAAATGACGGGTATTACAGACTA 1218
 QY 348 IleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGlu 367
 DB 1219 ACTGGGATGCGCACCATTAATCTCTGCAAGAGGTGGCTCCCGACGCTGCTCGAGAAC 1278
 QY 368 GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAla 387
 DB 1279 ATACACAGCAACTGCCAGGCCCAATATCAATGATTTTGGCCATTAGCAAACTAAGAAG 1338
 QY 388 AlaGlySerLeuProGlyTyrTyrIleLeuArgArgSerProGlnAspTyrAspSerPhe 407
 DB 1339 GCGGTAAACAGACTGGACTATATGTGTACATGACCCCTTAGGACTTCAACAAATAC 1398
 QY 408 LeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArg 427
 DB 1399 TTTCTGACCTTGTGTGTGAGGAGAAATGCTCATTTGAATATAAATCAACTGTTTGAATACG 1458
 QY 428 GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg 447
 DB 1459 AAGAATGAGATGGAGAAATACACCTCAGCGGGACTAAGAGGAACTTCAGTAACCTTAAG 1518
 QY 448 GluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeu 467
 DB 1519 GACCTTTTGAATGCTACCAGATGGAACCTGTGCGCTCAGACAGTATCATCTCCAGTTT 1578
 QY 468 ThrSerCysCysAlaProArgProLysGlyLysSerAsnLeuIleValValArgGly 487
 DB 1579 ACCAAATGCTGCCCCCAAGCCAAAGATAAATCAACCTTCTCGCTTCAGA----- 1632
 QY 488 CysAsnProAlaProAlaProGlyCysSerProSerCysCysAla-----LeuThr 504
 DB 1633 ACAATGCTATTTCTGATGTTTCAGATCTCCACACATTACAGAGGCATAAATGTGAAT 1692
 QY 505 GlnLeuSerPheHisThrIleProThrAspSerLeuGluThrHisGluAsnLeuGlyHis 524
 DB 1693 CAATGCTGTTTCAAAAATCAGGAATGAAGATTTAATATTTAATGAAGCTTGGGCGCAA 1752
 QY 525 GlySerPheThrLysIlePheArgGlySerArgGluValValAsp---GlyGluThr 543
 DB 1753 GGTACTTTTCAAAAAATTTTAAAGGTGAAGAGAGAGTGGAGATTTATGGTCAACTG 1812
 QY 544 HisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGluSer 563
 DB 1813 CAAAAACGGAAGTCTTTTGAAGTCTTAGATAAAGCACATAGGAACCTATTACAGAGTCT 1872
 QY 564 PheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuHis 583
 DB 1873 TTCTCGACACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1932
 QY 584 GlyValCysMetAlaGly---AspSerIleMetValGlnGluPheValTyrLeuGlyAla 602
 DB 1933 GGTGTCTGTCTGTGAGAGGAGAACATTTCTGTTCAAGAAATTTGAAATTTGAAAT 1992


```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 94..3480
US-08-665-574C-8

```

Alignment Scores:

Pred. No.:	1.01e-233	Length:	3629
Score:	2619.50	Matches:	537
Percent Similarity:	65.50%	Conservative:	190
Best Local Similarity:	48.38%	Mismatches:	350
Query Match:	44.70%	Indels:	33
DB:	3	Gaps:	15

US-09-397-967A-16 (1-1099) x US-08-665-574C-8 (1-3629)

QY	4	ProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSerSerGluAla	23
Db	145	CTGTACATCAAGAAATGGTGATATTTCTCGAAGTCTAATCTGTGAACGACAGATAGACCA	204
QY	24	GlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArgLeuSerPhe	43
Db	205	GTCCCTTCAAGTGATCTGTACCATCTCTGGCAAGCTGAAGAGATGATCTGAAGTTT	264
QY	44	SerpGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAlaCysGlyIle	63
Db	265	CCAAAGTGAGAGATGTTGACAGAAATTTGTGTGCTGCTCTCTAAAGCTCTGTGTAAT	324
QY	64	LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTyrPhePro	83
Db	325	ACGCTGTGTATCAATAATATTTGTGCGTTAATGAGTGAACCCGAAGGATCTGGTACCCA	384
QY	84	ProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg	103
Db	385	CCCATCATGTCTCCACATAGACGAGTCACACGAGCATGACATCTCTACAGGATAGG	444
QY	104	PheTyrPheProAspTyrPhe-----GlyLeuGluThrCysHisArgPheGlyLeuArg	121
Db	445	TTCTACTTCCCTCATTTGTACTGTAGTGGCAGCAGACCACTACAGATACCGAGTGTCC	504
QY	122	LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis	141
Db	505	CGTGGGCTGAAGCTCTCTGCTGTATGACTTTGTGTCATGCTTACCTTTTGTCTCAGTG	564
QY	142	ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGlu	161
Db	565	CGCATGATTTTGTTCACGGATGATAAAGTACCTGTGACTCATGAACCTCAGGAAGAG	624
QY	162	PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaargGluGlnAlaGlnArgPro	181
Db	625	TGCTTTGGATGGCGGTGTAGACATATGAGAATAGCTAAGAGAGAAACACGACTCCA	684
QY	182	GlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProSerLeuArgAspVal	201
Db	685	CTGGCTGTCATAACTCTGTTCAGCTACAGACATCTTACCAAGTGCCTGACGGCAG	744
QY	202	IleGlnGlyGlnAsnPheValThrArgArgIleArgArgThrValValLeuAlaLeu	221
Db	745	ATCCAAAGACTATCATTTTAAACCCGGAAGCGAATCAGGTACAGATTTCCGAGATTCAT	804
QY	222	LeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAsp	241
Db	805	CAGCAATTCAGTCAATGTAAAGCCACTGCCAGAACCTTAAACTTAAGTATCTTTATAAC	864
QY	242	LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal-----	256
Db	865	CTGGAACCCCTGCAGTCTGCCTCTACACAGACAGTTTGAAGTAAAGNAATCTGCAAGA	924
QY	257	GlyLeuProGlyAlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle	276
Db	925	GGT---CCTTCAGGTGAGAGATTTTGGCAACCATATATATACTCGAACCGGTGGAATT	981


```

; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 94...3480
US-08-946-994-8

```

Alignment Scores:

```

Pred. No.:      1,01e-233      Length:      3629
Score:          2619.50        Matches:      537
Percent Similarity: 65.50%      Conservative: 190
Best Local Similarity: 48.38%    Mismatches: 350
Query Match:      44.70%        Indels:      33
DB:              3             Gaps:       15

```

US-09-397-967a-16 (1-1099) x US-08-946-994-8 (1-3629)

```

QY 4 ProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSerSerGluAla 23
DB 145 CCTGTACATCAGATGGTGTATCTCTGGAAGTGTCTAATCTGTGAAGCAGATAGCCA 204
QY 24 GlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArgLeuSerPhe 43
DB 205 GTCCTTCAAGTGTATCTGTACCATCTCTTGGGCAAGCTGAAGGAGATCTGAAGTTT 264
QY 44 SerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAlaCysGlyIle 63
DB 265 CCAAGTGGAGAGTATGTTGCAGAGAANTTTGTGGCTGCTCTTAAGCTTGTGTATT 324
QY 64 LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPhePro 83
DB 325 ACGCTGTGTATCATATATGTTGCGTTAATAGTGAACCGAAAGGATCTGGTACCCA 384
QY 84 ProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg 103
DB 385 CCCAATCAATGCTTCCACATAGACAGTCAACCGGCATGACATCTCTACAGGATAAGG 444
QY 104 PheTyrPheProAspTyrPhePhePhePhePhePhePhePhePhePhePhePhePhe 121
DB 445 TTCTACTTCCCTCATCTGTACTGTAGTGGCAGCAGCAGCACTACAGATAGCGAGTGCC 504
QY 122 LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141
DB 505 CGTGGGCTGAAGTCTCTCTGATGACATGATGATGATGATGATGATGATGATGATGAT 564
QY 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGlu 161
DB 565 CGGCATGATTTTGTTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 624
QY 162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgPro 181
DB 625 TGTCTTGGGATGGCGGTGTAGACATGATGATGATGATGATGATGATGATGATGATGAT 684
QY 182 GlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal 201
DB 685 CTGCGCTGTCTATAACTCTGTCTGATCAAGACATCTTACCAAGTGGCTTCGAGCGAAG 744
QY 202 IleGlnGlnGlnAsnPheValThrArgArgIleArgThrValValLeuAlaLeu 221
DB 745 ATCCAGACTATCATTTTAAACCGGAGGAGATCAGGTACAGATTTCGACAGATTTCATT 804
QY 222 LeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAsp 241
DB 805 CAGCAATTTCAGTCAATTAAGCCACTGCCAGGACCTAAACCTTAAGTATCTTATTAAC 864
QY 242 LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal 256
DB 865 CTGGAACCCCTGCAGTCTGCTTCTACACAGAACAGTTTGAAGTAAAGAAATCTGCAAGA 924
QY 257 GlyLeuProGlyAlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle 276
DB 925 GGT---CCTTTCAGGTGAGGAGATTTTTCACCAACCATTAATAACTGGAAACCGGTGAATT 981

```

```

QY 277 ProTrpSerSer-----AsnAspGluLeuPheGlnThr 287
DB 982 CAGTGTCAAGAGGGAACAATAGGAAAGTGAACACTGACAGACAGACAGCTACAGTTA 1041
QY 288 PheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyPro 307
DB 1042 TATTGTGATTTCCCTGATATTATTGATGTCAATTAAGCAAGCAACACAG---GAATGC 1098
QY 308 AlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGlu 327
DB 1099 TCAATGAAGTAGAATTTGTAACCTCTCCATAAACAAGATGGTAAAGTTTTCGAGATAGA 1158
QY 328 PheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347
DB 1159 CTTAGCTCATTAATAAAGAGCCTTGTCTGTCTCATTAATTAAGCGGGTATTACAGACTA 1218
QY 348 IleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGlu 367
DB 1219 ACTGGGATGGCCACCATTAACCTCTGCAAGAGGTGGTCCCCAGCTGTGCTCGAGAAC 1278
QY 368 GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAla 387
DB 1279 ATACACAGCAACTGCCAGGCCCAATATCAATGGATTTTGCATTAGCAAACTAAAGAAG 1338
QY 388 AlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPhe 407
DB 1339 CGGGTAAACCACTGGGACTATATGTCTAGTGCAGCCCTTAAGGACTTCAACAAATAC 1398
QY 408 LeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArg 427
DB 1399 TTTCTGACCTTGTGTTGACGAGAAATGCTATTGAATATAAACCTGTGTGATTACG 1458
QY 428 GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg 447
DB 1459 AAGAATGAGATGGAGAATCAACCTCAGCGGCACTAAGAGGAACCTTCAGTAACCTTAAG 1518
QY 448 GluLeuLeuAlaAlaCysTyrPheAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeu 467
DB 1519 GACCTTTTGAATTTGCTCCAGATGGAACCTGTGCGCTCAGACAGTATCATCTCCAGTTT 1578
QY 468 ThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIleValValArgArgGly 487
DB 1579 ACCAAATGCTGCCCCCAAGCCAAAGATAAATAACAACTTCGCTTCAGACAG--- 1632
QY 488 CysAsnProAlaProAlaProGlyCysSerProSerCysCysAla-----LeuThr 504
DB 1633 ACAATGGTATTTTCTGATGTTTCAGATCTCACCACCAATTACAGAGGCATAATAATGTGA 1692
QY 505 GlnLeuSerPheHisThrIleProThrAspSerLeuGluThrHisGluAsnLeuGlyHis 524
DB 1693 CAATGTGTGTTTCAAAAATCAGGAATGAAGATTTAATATTTAATGAAGTCTTGGCCAA 1752
QY 525 GlySerPheThrLysIlePheArgGlySerArgGluValValAsp---GlyGluThr 543
DB 1753 GGTAATTTTCAAAAATTTTAAAGGTGAAGAGAGAGTGGAGATTTATGTTCAACTG 1812
QY 544 HisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGluSer 563
DB 1813 CAAAAACGGAAGTCTTTTGAAGTCTTAGATAAAGCACATAGGAACCTATTACAGAGTCT 1872
QY 564 PheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuHis 583
DB 1873 TTTCTGAGACAGACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1932
QY 584 GlyValCysMetAlaGly---AspSerIleMetValGlnGluPheValTyrLeuGlyAla 602
DB 1933 GGTGTCTGTCTGTGGAGGAGAACATTTCTGTTCAAGAAATTTGTAAATTTGGATCA 1992
QY 603 IleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLysLeuGlnVal 622
DB 1993 CTGGATACATACCTGAAGAGAACAAATTTCCATAAATATATATATGGAACCTTGAGTG 2052
QY 623 ThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsn 642

```

Db	2053	GCTAAGCAGTGGCAGTGGCCATGATTTCTAGAGAAATCCCTATTATCGGAAT	2112	Db	3124	AAGGAGCCAGGGGAAAGCCCATATTTCTGGTACGACCCCTGAATCTTTCAGGAGCAAG	3183
Qy	643	ValSerAlaArgLysValLeuLeuAlaArgGluGlyGly-AspGlyAsnProPro	660	Qy	999	PheSerArgGlnSerAspValTrpSerPheGlyValValLeuLeuPheThrTyr	1018
Db	2113	GTGTGTCTAAATAATCTCTGCTTATCAGAGAAGAAGCAGGAGAACGGGAAACCCACCT	2172	Db	3184	TTTTCTGTGGCTCAGATGTGTGGAGCTTGGAGTGGTCTATACGAACCTTTTCATAC	3243
Qy	661	PheLeuLysLeuSerAspProGlyValSerProThrValLeuLeuSerLeuMetLeuThr	680	Qy	1019	CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGlu	1038
Db	2173	TTTCATCAAACTAGTATCTCTGGCATTTAGCATTTACAGTTCTTACCGAAGGACATTTCTAG	2232	Db	3244	ATCGAGAAGAGTAAAGTCCACCCCTGGGAATTTATGCGAATGATTGGCAATGATAACAA	3303
Qy	681	AspArgLysProThrValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu	700	Qy	1039	GlyProProLeuCys-ArgLeuLeuGluLeuAlaGluGlyArgArgLeuProPro	1057
Db	2233	GAGAAATACCATGGTCTCTGATGATCATTTAGAAATCTTAAATCTTCAATCTGGCA	2292	Db	3304	GGGCAATGATTTGTTCATTTGATAGAGTACTGAGAGCAACGGAAGATTGCCAAG	3363
Qy	701	AlaAspLysTrpGlyPheGlyAlaThrTrpGluValPheGlnArgGlyProAlaHis	720	Qy	1058	ProProThrCysProThrGluValGlnLeuMetGlnLeuCysTrpAlaProGluPro	1077
Db	2293	ACAGACAGTGGAGCTTCGGACCACTCTGTGGAGATCTGCAGTGGAGAGATAAGCCC	2352	Db	3364	CCAGAGAGTGGCAGATGAGATTTATGTGATCATGACAGAGTGTCTGGAACCAATGTG	3423
Qy	721	IleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeuPro	740	Qy	1078	HisAspArgProAlaPheAlaThrLeuSer	1087
Db	2353	CTGAGTCTCTGATTTCTCAAAAGAAAGCTGCAATCTTGAAGATAAGCATCAGCTTCT	2412	Db	3424	AGCCAGCGTCCCTCTTCAGGACCTTTTCG	3453
Qy	741	AlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly	760				
Db	2413	GCACCAAGTGGACAGATTAGCAACCTTAAATTAATTTGATGAGTATGAGCCAGAT	2472				
Qy	761	ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr	780				
Db	2473	TTTCAGGCTCTTTTCAGAGCTGTCTCATCGGTGATCTTAAACAGCCTGTTTATCTCCAGAT	2532				
Qy	781	GluLeuLeuSerAspProThrProGlyIleProSerProArg-AspGluLeuCysVal	799				
Db	2533	GAACCTACTAACGAA--AATGACATCTCTACCAACATGAGATAGTGGCCCTAGGGTTT	2589				
Qy	800	AlaGlyValGlnLeuThrAlaCysGlnAspAlaIlePheGluGluArgHisLeuLys	819				
Db	2590	TTCTGGTGTCT--TTTGAAGCAGGACCCCTACACAGTTTGAAGAGACACTTGAAG	2643				
Qy	820	TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspPro	839				
Db	2644	TTTCTACAGCAGCTTGGCAAGGTAACTTCGGAGTGTGGAGATGTGCCGCTATGACCCG	2703				
Qy	840	LeuGlyAspAsnThrGlyProLeuValAlaVallyGlnLeuGlnHisSerValProAsp	859				
Db	2704	CTGCAGGACAACTGCGGAGTGTGCTGTGAAGAACTCCAGCACAGCCTGAAGAG	2763				
Qy	860	GlnGlnArgAspPheGlnArgGluIleGlnLeuLysAlaLeuHisSerAspPheIle	879				
Db	2764	CACCTCGAGACTTTGAGGGAGATCGATCTCGAAATCTTCGACGATGACAAATC	2823				
Qy	880	VallyTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu	899				
Db	2824	GTCAAGTACAAGGAGTGTCTACAGTGGGTGCGGCTCGGCAACCTTAAGATTAAATTTATGAA	2883				
Qy	900	TyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg--GlyLeuHisThr	918				
Db	2884	TATTTACCATATGGAAGTTTACGAGACTTCTCAAAATAAATAAGAACGGATAGATCAC	2943				
Qy	919	AspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArg	938				
Db	2944	AAAAATCTTCTCAATACATCTCGATATGCAAGGCGCATGGAATATCTGTTGATACAAA	3003				
Qy	939	ArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisVal	958				
Db	3004	AGGTATATCCAGGAGCTGGCAACAGGAACATTTGTGTGAAATGAGAACAGGGTT	3063				
Qy	959	LysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVal	978				
Db	3064	AAAAATAGGAGACTTCTGGATTAAACCAAGTCTTGGCGAGGCAAGAAATATCAAAAGTA	3123				
Qy	979	ArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAsnIle	998				

Alignment Scores: 2.9e-231 Length: 4078
Pred. No.: 2594.00 Matches: 531
Score:

RESULT 13

US-09-016-434-297
; Sequence 297, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 297:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4078 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PLACNOB01
; CLONE: 179527
; US-09-016-434-297

Percent Similarity:	65.94%	Conservative:	193
Best Local Similarity:	48.36%	Mismatches:	343
Query Match:	44.27%	Indels:	32
DB:	4	Gaps:	14
US-09-397-967a-16 (1-1099) x US-09-016-434-297 (1-4078)			
Qy	20	SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ArgGlyProGlyProGlyProGly 39	
Db	447	TCAGTCTTCAGGT-----GTATCTTTACCATTCCTTGGAAATCTGAGGCAGA 497	
Qy	39	naArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaIy 59	
Db	498	TTATCTGACCTTTCATCTGGGAGTATGTTGGAGAAGAAATCTGTATTGCTGCTCTAA 557	
Qy	59	sAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPhe 79	
Db	558	AGCTTGTGTATCATCACCCTGTGTATCAATAATGTTGCTTTAATGAGTGAACAGAAAG 617	
Qy	79	rCysTyrPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa 99	
Db	618	GATCTGGTATCCACCAACCATGCTCTTCATATAGATGATCAACAGGCATATGTACT 677	
Qy	99	lTyrArgLeuArgPheTyrPheProAspTyrPhe-----GlyLeuGluThrCysHisAr 117	
Db	678	CTACAGAATAAGATTTTACTTTCCTCG-TGGTATTCAGTGGCAGCAACAGAGCCTATCG 736	
Qy	117	gPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137	
Db	737	GCATGGATATCTCAGGTGCTGAAGCTCCTCTCTTGATGACTTTGTCATGTCCTTACT 796	
Qy	137	uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157	
Db	797	CTTTGCTCAGTGGCGCATGATTTTGTGCATGGATGGATAAAGTACCTGTGACTCATGA 856	
Qy	157	sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl 177	
Db	857	AACACAGGAAGATGTCTTGGGATGACAGTGTAGATATGATGATGAGATAGCCAAAGAAA 916	
Qy	177	naIaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSe 197	
Db	917	CGATCAAAACCCACTGGCCATCTTAACCTATCAGCTTACAGACATTCCTTACCACAATG 976	
Qy	197	rLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgThrVa 217	
Db	977	TATTCGAGCAAGATCCAAAGACTATCATATTTTGCAAGGAGCGAATAAGGTACAGATT 1036	
Qy	217	lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLy 237	
Db	1037	TCGCAGATTTTATTCAGCAATTCAGCCAAATGCAAGCCCACTGCCAGAAACTTGAAACTTA 1096	
Qy	237	sTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGl 257	
Db	1097	GTATCTTATAAATCTGGAACCTCTGCACTCTGCGCTCTTACACAGAGAAATTTGAAGTAA 1156	
Qy	257	yLeuProGlyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAs 273	
Db	1157	AGAACCTTGAAGTGTCTCTCAGGTGAGGAGATTTTGCAACCATTTATATTAACCTGAAA 1216	
Qy	273	pAsnGlyIleProTyrSerSer-----AsnAspGluLe 284	
Db	1217	CGGTGGAATTCAGTGGTCAAGAGGGAACATAAAGAAAGTGGAGACCTGACAGAACAGGA 1276	
Qy	284	uPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProAr 304	
Db	1277	TTTACATTTATTTGCGGATTTCTTAATATTTATTTGATGTCTAGTATTAAAGCAAGCAACCA 1336	
Qy	304	gValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLe 324	
Db	1337	AGAGGGT---TCRAATGAAGCCAGTTGTAACTATCCATAAGCAAGATGGTAAATAATCT 1393	
Qy	324	uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTy 344	
Db			
Db	1394	GGAAATTGAACCTTAGCTCATTAAGGGAAGCTTTGCTTTCGTGTCTATTAATTGATGGATA 1453	
Qy	344	rPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLe 364	
Db	1454	TTATAGATTAACTGCAGATGCACATCATCTCTGTAAAGAAGTAGCACCTCCAGCCGT 1513	
Qy	364	uLeuGluGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLy 384	
Db	1514	GCTTGAATAATATACAAAGCAACTGTCTATGGCCCAATTTTCGATGGATTTTGGCCATTAGTAA 1573	
Qy	384	sLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGluAspTy 404	
Db	1574	ACTGAAGAAAGCAGGTAATCAGACTGGAGTGTATGTACTTCGATCGACTCTTAAGGACTT 1633	
Qy	404	rAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCy 424	
Db	1634	TAATAAATATTTTGTGACTTTTGTCTGTCGAGGAGAAAATGTCTTGAATATATAACACTG 1693	
Qy	424	sLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAr 444	
Db	1694	TTTGATTTCAAAAATGAGATGAGAGTACAACCTCAGTGGGCAAAAGAAAGACTTCAG 1753	
Qy	444	gSerLeuArgGluLeuLeuAlaAlaCysTyrAsnSerGlyLeuArgValAspGlyAlaAl 464	
Db	1754	CAGTCTTAAAGATCTTTTGAATTTTACAGATGAAACTGTTCCGCTCAGACAATATATAT 1813	
Qy	464	aLeuTyrLeuThrSerCysCysAlaProArgProLysGlyLysSerAsnLeuIleValVa 484	
Db	1814	TTTCCAGTTTACTAATGCTGTCCCAAGCAAAAGATAAATCAAACTTCTAGTCTT 1873	
Qy	484	lArg---ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysCysAlaLe 503	
Db	1874	CAGAACGATGCTGTTCTGATGTACCACTCCACCAACATTACAGAGGCTTACTCATAT 1933	
Qy	503	uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTyrHisGluAsnLeuGl 523	
Db	1934	GAACCAATGCTGTTTTCACAAATCAGAAATGAAGATTGTATATTATTAAGAGCCTTGG 1993	
Qy	523	yHisGlySerPheThrLysIlePheArgGlySerArgArgGluValValAsp---GlyGl 542	
Db	1994	CAAAGGCACTTTTACAAGATTTTAAAGGCTGACGAAGAGAGAGTAGGAGACTACGGTCA 2053	
Qy	542	uThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGl 562	
Db	2054	ACTGCATGAACAGAGTCTTTTAAAGTCTGATGATAAGCAACACACAGGAACTATTCAGA 2113	
Qy	562	uSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLe 582	
Db	2114	GTCTTTCTTTGAGCAGCAAGTATGATGAGCAAGCTTTTCACAGCATTTGGTTTTAAA 2173	
Qy	582	uHisGlyValCysMetAlaGlyAsp---SerIleMetValGlnGluPheValTyrLeuGl 601	
Db	2174	TTATGGAGTATGTCTGTGAGCAGCAGATAATTTCTGGTTTCTGAGGATTTGTAAATTTGG 2233	
Qy	601	yAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrLysLeuGl 621	
Db	2234	ATCATTAGTACATATCTGAAAGAAATAAAAAATTTGATAAATATATATTATGGAACACTGA 2293	
Qy	621	nValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGl 641	
Db	2294	AGTTGCTAAACAGTTGGCATGGGCCCATGTCATTTCTAGAGAAACACACCTTTATTCTAAG 2353	
Qy	641	yAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGly-----AspGlyAsnPr 659	
Db	2354	GAATGTATGTGCCAAAAATATCTGCTTATCAGAGAAGAGACAGGAGACAGGAAATCC 2413	
Qy	659	oProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLe 679	
Db	2414	TCCTTTTCATCAAACTTAGTGATCTCGCATATTAGTATTACAGTTTTCGCAAGACATCTCT 2473	
Qy	679	uThrAspArgIleProTyrValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLe 699	
Db	2474	TCAGGAGAGATACCATGGGTACCACTGNATGCATTTGAAATCTCTAAATTTAAATTT 2533	

QY	699	uCluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAl	719	QY	1056	pProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGl	1076
Db	2534	GCACACACAAATGGAGTTTGGTACCACTTTGGGGAATCTCGAGTGGAGAGATAA	2593	Db	3605	AAGACACAGATGATCCAGATGATCTATATGATCATGACAGATGCTGAACAATAA	3664
QY	719	aHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnIle	739	QY	1076	uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu	1093
Db	2594	ACCTCTAAGTCTCTGGATTCTCAAGAAAGCTACAAATTTTATGAAGATAGGCATCAGCT	2653	Db	3665	TGTAATCAACGCCCTCTCTTAGGATCTAGTCTCTTGAGTGGATCAATAA	3716
QY	739	uProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPr	759	RESULT 14			
Db	2654	TCCTGCACCAAGTGGCAGAAATTAGCAACCTTATAATAATTGTATGGATTATGAACC	2713	US-08-446-038B-2			
QY	759	oGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs	779	Sequence 2, Application US/08446038B			
Db	2714	AGATTTTCAGGCTTCTTTTCAGAGCCCATCATACAGATCTTAACAGTTTGTACTCCAGA	2773	Patent No. 5658791			
QY	779	pTyrGluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCy	798	GENERAL INFORMATION:			
Db	2774	TTATGAACCTATTAAACAGAA---AATGACATGTTTACCAATATCAGGATAGGTGCCTGGG	2830	APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;			
QY	798	sValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisIle	818	APPLICANT: Harpur, Ailsa			
Db	2831	GTITTTCTGTGTGCC-----TTTGAAGACCGGGATCCTCACACAGTTTGAAGAGAGACATTT	2884	TITLE OF INVENTION: No. 5658791el Protein Tyrosine Kinase			
QY	818	uLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAs	838	NUMBER OF SEQUENCES: 23			
Db	2885	GAAATTTCTACGCAACTTGGCAAGGTAAATTTGGAGGTGAGATGTCGGGTATGA	2944	CORRESPONDENCE ADDRESS:			
QY	838	pProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPr	858	ADDRESS: 805 Third Avenue			
Db	2945	CCCTCTACAGGACACACATGGCGAGTGGTCTGTAAAAAGCTTCAGCATAGTACTGA	3004	CITY: New York City			
QY	858	oAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPh	878	STATE: New York			
Db	3005	AGAGCACCTAAGAGACTTTGAAAGGGAATTTGAAATCCCTGAAATCCCTACAGCATGACAA	3064	COUNTRY: USA			
QY	878	eIleValLysTyrArgLysValSerTyrGlyProGlyArgGlnSerLeuArgValLe	898	ZIP: 10022			
Db	3065	CATTGTAAGTACCAAGGGATGTCTACAGTGTGGTGGCGTAACTAAATAATATAT	3124	COMPUTER READABLE FORM:			
QY	898	tGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHi	917	MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage			
Db	3125	GGATATTTTACCATATGGAAGTTTACGAGACTATCTTCAAAAAACATAAAGACGGATAGA	3184	COMPUTER: IBM PS/2			
QY	917	sThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAl	937	OPERATING SYSTEM: PC-DOS			
Db	3185	TCACATAAAACTTCTGCAGTACACATCTCAGATATGCAAGGGTATGGAGTATCTTGTGAC	3244	SOFTWARE: Wordperfect			
QY	937	aArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHi	957	CURRENT APPLICATION DATA:			
Db	3245	AAAAGGTATATCCACAGGATCTGGACAGAGAAATATATTTGGTGGAGACGAGACAG	3304	APPLICATION NUMBER: US/08/446,038B			
QY	957	sValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrIleVa	977	FILING DATE: 19-MAY-1995			
Db	3305	AGTTAAATTTGGRATTTGGTTTAAACCAAGTCTGCCACACACAAAGAACTACTATAA	3364	CLASSIFICATION: 424			
QY	977	lValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAs	997	PRIOR APPLICATION NUMBER: 08/064,067			
Db	3365	AGTAAAGAAACCTGGTGAAGTCCCATATTCTGGTATGCTCCAGAAATCATGACAGAGAG	3424	FILING DATE: 30-Jun-1993			
QY	997	nIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheTh	1017	APPLICATION NUMBER: PCT/US91/08889			
Db	3425	CAAGTTTCTGTGGCTTCAGATGTTTGGAGCTTTGGAGTGGTCTGTATGAACCTTTTAC	3484	FILING DATE: 26-No. 5658791-1991			
QY	1017	rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluAr	1037	ATTORNEY/AGENT INFORMATION:			
Db	3485	ATACATTGAGAAGAGTAAAGTCCACAGCGGAATTTATGCGTATGATTTGGCAATGACAA	3544	NAME: Hanson, No. 5658791man D.			
QY	1037	gGluGlyProProLeuCys---ArgLeuLeuGluLeuAlaGluGlyArgArgLeuPr	1056	REGISTRATION NUMBER: 30,946			
Db	3545	ACAAGGACAGATGATCGTTCCTTTCATGATGAAGAACTTTTGAAGAAATAATGAAGATTACC	3604	REFERENCE/DOCKET NUMBER: LUD 5244			
				TELEPHONE: 212-688-9200			
				TELEFAX: 212-838-3884			
				INFORMATION FOR SEQ ID NO: 2:			
				SEQUENCE CHARACTERISTICS:			
				LENGTH: 3495 base pairs			
				TYPE: nucleic acid			
				STRANDEDNESS: single			
				TOPOLOGY: linear			
				MOLECULE TYPE: nucleic acid			
				US-08-446-038B-2			
				Alignment Scores:			
				Pred. No.: 1,44e-209			
				Score: 2360.00			
				Percent Similarity: 65.89%			
				Best Local Similarity: 49.34%			
				Query Match: 40.27%			
				DB: 1			
				Length: 3495			
				Matches: 489			
				Conservative: 164			
				Mismatch: 300			
				Indels: 38			
				Gaps: 15			
				US-09-397-967A-16 (1-1099) x US-08-446-038B-2 (1-3495)			
QY	128	IleLeuAspLeuHisValLeuGluHisValLeuPheAlaGlnHisArgSerAspLeuValSer	147	QY	128	IleLeuAspLeuHisValLeuGluHisValLeuPheAlaGlnHisArgSerAspLeuValSer	147
Db	1	CTGCTTATGACCTTTGTTCATGCTCTTACCTTCCCTCAGTGGCGCATGATTTTGTTCAC	60	Db	1	CTGCTTATGACCTTTGTTCATGCTCTTACCTTCCCTCAGTGGCGCATGATTTTGTTCAC	60

148 GlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGluPheLeuSerLeuAlaVal 167
1129 GTTCAGCTCTCACCACCAATTACAGAGCATATAATGTGAATCAATGGTGTTCACCAA 1188
511 IleProThrAspSerLeuGluTriPheHisGluLeuGlyHisGlySerPheThrLysIle 530
1189 ATCAGGAATGAAGATTTCATATTAATGAAGCCTTGCCAGGACACTTTTACAAAATA 1248
531 PheArgGlySerArgArgGluValValAsp---GlyGluThrHisAspSerGluValLeu 549
1249 TTTAAAGGTGAAGAGAGAGAGTTGGAGATTATGTGTCAGTCGACGACGAAACCGAAGTCTT 1308
550 LeuLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaIle 569
1309 TTGAAGTCCTAGATTAAGACATAGAAACTATTTCAGAGTCTTTCTTTGAAGCAGCAAGC 1368
570 LeuMetSerGlnValSerProHisLeuValLeuLeuHisGlyValCysMetAlaGly 589
1369 ATGATGAGTCAGCTTCTCACAGCAATTTGGTTTGAATATGAGATATGTCTGTGGA 1428
590 ---AspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArg 608
1429 GAGGAGAACATTTGGTTTCAGAGATTGTAAATTTGGATTCATGATACATACCTGAAG 1488
609 LysArgGlyHisLeuValSerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyr 628
1489 AAGAACAAAATCTATAAATATATTATGGAACCTTGGAGTGGGAGCAGTGGCATG 1548
629 AlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysVal 648
1549 GCCATGCATCTCTCGAAGAAAATCCCTTATTTCATGGAAATGTGTGCTTAAATAATC 1608
649 LeuLeuAlaArgGluGlyGly---AspGlyAsnProPropheIleLysLeuSerAsp 666
1609 CTGCTTATCAGAGAAGAAGACAGAGAACCGGGAACCCACCTTTCATCAAACTTAGTGAT 1668
667 ProGlyValSerProThrValLeu-----SerLeuGluMetLeu 679
1669 CCTGGCAATTAGCATTTACAGTTCTACCGAAGGACATTTCTCTCTGTTTCCAAAGTCTT 1728
680 ThrArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeu 699
1729 CAGGAGAGATACCATGGTACACCTGATGATGAGATCCTAAATCTTAAATCTAAGCTG 1788
700 GluAlaAspLysTrpGlyPheGlyAlaThrTrpGluValPheGlnArgGlyProAla 719
1789 GCAACAGACAAGTGGAGCTTCGGGACCACTCTGTGGAGATCTCGAGTGGAGGAGATAAG 1848
720 HistLeuThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeu 739
1849 CCCCTGAGTCTCTGGATTCTCAAGAAAGCTGAGTTCTTATGAAGATAAGCATCAGCTT 1908
740 ProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPro 759
1909 CCTGCACCAAGTGGACAGATTGGCAACCTTATAAATATTCATGAGTGGATGAGCCA 1968
760 GlyArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAsp 779
1969 GATTTCAGGCTGCTTTCAGAGCTGTCATCGGTGATCTTAAACAGCCTGTTTACTCCAGAT 2028
780 TyrGluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCys 798
2029 TATGAACACTCAACAGAA---AATGACATGCTACCAACATGAGATAGTGGCTTAGGG 2085
799 ValAlaGlyValAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeu 818
2086 TTTTCTGGTCT-----TTTGAAGACAGGACCCCTACACAGTTTGAAGAGAGACACTT 2139
819 LysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAsp 838
2140 AAGTTTCTACAGCAGCTTGGCAAGGTAATCTCGGAGTGTGGAGATGTGCCCTATGAC 2199
839 ProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPro 858

QY 494 ProGlyCysSerProSerCysCysAla-----LeuThrGlnLeuSerPheHisThr 510
Db 1129 GTTCAGCTCTCACCACCAATTACAGAGCATATAATGTGAATCAATGGTGTTCACCAA 1188
QY 511 IleProThrAspSerLeuGluTriPheHisGluLeuGlyHisGlySerPheThrLysIle 530
Db 1189 ATCAGGAATGAAGATTTCATATTAATGAAGCCTTGCCAGGACACTTTTACAAAATA 1248
QY 531 PheArgGlySerArgArgGluValValAsp---GlyGluThrHisAspSerGluValLeu 549
Db 1249 TTTAAAGGTGAAGAGAGAGTTGGAGATTATGTGTCAGTCGACGACGAAACCGAAGTCTT 1308
QY 550 LeuLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaIle 569
Db 1309 TTGAAGTCCTAGATTAAGACATAGAAACTATTTCAGAGTCTTTCTTTGAAGCAGCAAGC 1368
QY 570 LeuMetSerGlnValSerProHisLeuValLeuLeuHisGlyValCysMetAlaGly 589
Db 1369 ATGATGAGTCAGCTTCTCACAGCAATTTGGTTTGAATATGAGATATGTCTGTGGA 1428
QY 590 ---AspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArg 608
Db 1429 GAGGAGAACATTTGGTTTCAGAGATTGTAAATTTGGATTCATGATACATACCTGAAG 1488
QY 609 LysArgGlyHisLeuValSerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyr 628
Db 1489 AAGAACAAAATCTATAAATATATTATGGAACCTTGGAGTGGGAGCAGTGGCATG 1548
QY 629 AlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysVal 648
Db 1549 GCCATGCATCTCTCGAAGAAAATCCCTTATTTCATGGAAATGTGTGCTTAAATAATC 1608
QY 649 LeuLeuAlaArgGluGlyGly---AspGlyAsnProPropheIleLysLeuSerAsp 666
Db 1609 CTGCTTATCAGAGAAGAAGACAGAGAACCGGGAACCCACCTTTCATCAAACTTAGTGAT 1668
QY 667 ProGlyValSerProThrValLeu-----SerLeuGluMetLeu 679
Db 1669 CCTGGCAATTAGCATTTACAGTTCTACCGAAGGACATTTCTCTCTGTTTCCAAAGTCTT 1728
QY 680 ThrArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeu 699
Db 1729 CAGGAGAGATACCATGGTACACCTGATGATGAGATCCTAAATCTTAAATCTAAGCTG 1788
QY 700 GluAlaAspLysTrpGlyPheGlyAlaThrTrpGluValPheGlnArgGlyProAla 719
Db 1789 GCAACAGACAAGTGGAGCTTCGGGACCACTCTGTGGAGATCTCGAGTGGAGGAGATAAG 1848
QY 720 HistLeuThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeu 739
Db 1849 CCCCTGAGTCTCTGGATTCTCAAGAAAGCTGAGTTCTTATGAAGATAAGCATCAGCTT 1908
QY 740 ProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPro 759
Db 1909 CCTGCACCAAGTGGACAGATTGGCAACCTTATAAATATTCATGAGTGGATGAGCCA 1968
QY 760 GlyArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAsp 779
Db 1969 GATTTCAGGCTGCTTTCAGAGCTGTCATCGGTGATCTTAAACAGCCTGTTTACTCCAGAT 2028
QY 780 TyrGluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCys 798
Db 2029 TATGAACACTCAACAGAA---AATGACATGCTACCAACATGAGATAGTGGCTTAGGG 2085
QY 799 ValAlaGlyValAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeu 818
Db 2086 TTTTCTGGTCT-----TTTGAAGACAGGACCCCTACACAGTTTGAAGAGAGACACTT 2139
QY 819 LysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAsp 838
Db 2140 AAGTTTCTACAGCAGCTTGGCAAGGTAATCTCGGAGTGTGGAGATGTGCCCTATGAC 2199
QY 839 ProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPro 858

	DB	418	GAGATTTTTCACCACTTATTATAACTGGAAACCGTGGAATTCAGTGTGCTCAGAGGGGAAA	477
	QY	281	-----AsnAaspGluLeuPheGlnThrPheCysAaspPheProGlu	293
	DB	478	CATAAGGAAAGTGACACTGACAGAACGAGCGTACAGTTATATGTGATTTCCCTCAT	537
	QY	294	IleValaspValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeu	313
	DB	538	ATTATGTGATGTCAAGTAATTAAGCAAGCAAATCAG--GAATGCTCAACTGAAAGTAGAGTT	594
	QY	314	ValThrValThrArgMetAaspGlyHisIleLeuGluAlaGluPheProGlyLeuProGlu	333
	DB	595	GTGACCCTGCCAAGCAGGAGCGGAAGGCTCTTGGAAATAGAGAATAGACTTAGCTCTATTAAAGAA	654
	QY	334	AlaLeuSerPheValAlaLeuValAaspGlyTyrPheArgLeuIleCysAaspSerArgHis	353
	DB	655	GCTTGTTCATTCGTCTATTAAATGACGGGTATTACAGCTTAAGTGGGATGCACACCAT	714
	QY	354	TyrPheCysLysGluValAlaProProArgLeuLeuGluGluAlaAaspValCysHis	373
	DB	715	TACCTCTGCAAGAGGTGGCTCCCCCAGCTGTGTTTCGAGACAATACACAGCAACTGCCAC	774
	QY	374	GlyProIleThrLeuAaspPheAlaIleHisIlysIleuIysAlaAlaGlySerLeuProGly	393
	DB	775	GGCCCAATTTCAATGGATTTTGGCCATCAGCAAACTTAAAGAAGGAGGAGAAACACGACTGGA	834
	QY	394	ThrTyrlleLeuArgArgSerProGlnAaspTyrAaspSerPheLeuThrAlaCysVal	413
	DB	835	CTGTATGCTACTTCGATGTAGCCCTAAGGACTTCAACAATACTTCTTGACCTTTGCCGCTT	894
	QY	414	GlnThrProIleuGlyProAaspTyrLyseGlyCysLeuIleArgGlnAaspProSerGlyAla	433
	DB	895	GAGCGAGAAATGTTATTGAATAATAACACTGTTTGTGATTACAAAGAAATGAGAATGGAGAG	954
	QY	434	PheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuAlaAlaCys	453
	DB	955	TACAACCTCAGTGGGACTTAAGAGAACTTCAGTAGCTCTTAAGGACCTTTTGAATTGCTTAC	1014
	QY	454	TrpAsnSerGlyLeuArgValAaspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaPro	473
	DB	1015	CAGATGGAAACTGTGCGCTCAGACAGTATCATCTCCAGTTCACCAATGCTGTCCTCCA	1074
	QY	474	ArgProLysGluLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAla	493
	DB	1075	AAGCCGAAAGATAAATCAAACTTCTGTCTTCAGA----ACAAATGGTGTTCCTGAT	1128
	QY	494	ProGlyCysSerProSerCysCysAla-----LeuThrGlnLeuSerPheHisThr	510
	DB	1129	GTTCCAGCTCTCACCAACATTACAGAGCATTAATAATGTGAATCAATGGTGTTCACAAA	1188
	QY	511	IleProThrAaspSerLeuGluTyrHisIleGluAsnLeuGlyHisGlySerPheThrLysile	530
	DB	1189	ATCAGGAATCAAGATTGATATTAAATGAAGCCCTGGCCACAGCATTTTTACAAAAATA	1248
	QY	531	PheArgGlySerArgArgGluValValAasp---GlyGluThrHisAaspSerGluValLeu	549
	DB	1249	TTTAAAGGTGTAAGAAGAGAAGTTGGAGATTATGTCAGCTGCAGCAAAACCGAAGTCTT	1308
	QY	550	LeuLysValMetAaspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSer	569
	DB	1309	TTGAAGTCTAGATAAAGCACATAGAAACCTATTTCAGAGTCTTTTGAAGCAGCAAGC	1368
	QY	570	LeuMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGly	589
	DB	1369	ATGATGAGTCAGCTTCTCACAGCATTTGGTTTTGNAATTAGAGTATGTGCTGTGGA	1428
	QY	590	---AaspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAaspMetTyrLeuArg	608
	DB	1429	GAGGAGAACAATTTTGGTTCAAGAGTTTGTAAAAATTGGATCCTGGATACATACCTGAAG	1488
	QY	609	LysArgGlyHisLeuValSerAlaSerTyrLysLeuGlnValThrLysGlnLeuAlaTyr	628

QY 978 ValArgGluProGlyGlnSerProIlePheThrTyrAlaProGluSerLeuSerAspAsn 997
 Db 2620 GTAAAGGAGCCAGGGGAAAGCCCATATTCGTACGCACCTGAATCCTTGACGGAGAGC 2679
 QY 998 IlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThr 1017
 Db 2680 AAGTTTTCGTGGCCTCAGATGTGTGGAGCTTTGGAGTGGTTCTATACGAACCTTTTCACA 2739
 QY 1018 TyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArg 1037
 Db 2740 TACATCGAAGAGTAAAGTCCACCCGCGGAATTTATCGGAATGATTGGCAATGATAAA 2799
 QY 1038 GluGlyProProLeuCys---ArgLeuLeuGluLeuAlaGluGlyArgArgLeuPro 1056
 Db 2800 CAAGGGCAAAATGATTGTGTTCATTGTGATAGAGCTACTGAGAGCAACGGAAAGATTGCCA 2859
 QY 1057 ProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGlu 1076
 Db 2860 AGGCCAGAGGATGCCAGATGAGATTTATGTGATCATGCACAGAGTGTCTGGAAACAACAT 2919
 QY 1077 ProHisAspArgProAlaPheAlaThrLeuSer 1087
 Db 2920 GTGAGCCAGCGTCCCTCCTTCAGGGACCTTTCC 2952

Search completed: February 4, 2004, 06:28:04
 Job time : 295.373 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2004, 06:28:12 ; Search time 1087.04 Seconds
(without alignments)
3724.149 Million cell updates/sec

Title: US-09-397-967A-16

Perfect score: 5860

Sequence: 1 MAPSEETPLIPQSCSLSS.....RPAPATLSPQLDPLWRGPG 1099

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09397967/runat_03022004_175633_28650/app_query.fasta_1.1486
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HRAPIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09397967@cgn1.1.383@runat_03022004_175633_28650
-NCPU=6 -ICPU=3 -NO.WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

	1	4713.5	80.4	5118	13	US-09-814-353-19356
2	4597	78.4	3620	12	US-10-305-720-1055	Sequence 1356, A
3	2636.5	45.0	3731	12	US-10-191-803-200	Sequence 1055, Ap
4	2594	44.3	4078	12	US-10-305-720-297	Sequence 200, App
5	2467	42.1	7126	12	US-10-062-674-2214	Sequence 297, App
6	1927	32.9	5072	12	US-10-295-027-697	Sequence 2214, Ap
7	1927	32.9	5073	15	US-10-084-817-230	Sequence 697, App
8	1908.5	32.6	4191	11	US-09-962-854A-1	Sequence 230, App
9	1904	32.5	3541	10	US-09-880-107-2379	Sequence 1, Appli
10	1904	32.5	3541	13	US-09-873-367C-163	Sequence 2379, Ap
11	1842	31.4	3538	11	US-09-962-854A-2	Sequence 163, App
12	1817.5	31.0	4080	12	US-10-305-720-1353	Sequence 2, Appli
13	1446.5	24.7	5228	12	US-10-062-674-2215	Sequence 1353, Ap
14	1281.5	21.9	2597	12	US-10-205-219-2	Sequence 2215, Ap
15	746	12.2	2327	12	US-10-131-410-22	Sequence 2, Appli
16	713	12.2	2338	9	US-09-925-302-90	Sequence 22, Appl
17	675.5	11.5	1179	15	US-10-211-088-131	Sequence 90, Appl
18	609.5	10.4	2027	10	US-09-771-161A-29	Sequence 131, App
19	483.5	8.3	975	11	US-09-962-854A-5	Sequence 29, Appl
20	482.5	8.2	753	9	US-09-910-943-636	Sequence 5, Appli
21	482	8.2	3845	13	US-10-007-926A-110	Sequence 636, App
22	482	8.2	3845	13	US-10-101-510-4	Sequence 110, App
23	482	8.2	3845	13	US-10-021-660-34	Sequence 4, Appli
24	482	8.2	3914	14	US-10-044-090-148	Sequence 34, Appl
25	475	8.1	3765	13	US-10-207-498-5	Sequence 148, App
26	475	8.1	3768	9	US-09-811-123-8	Sequence 5, Appli
27	475	8.1	3768	9	US-09-811-115-2	Sequence 8, Appli
28	475	8.1	4530	10	US-09-877-177-11	Sequence 2, Appli
29	475	8.1	4530	12	US-10-159-563-208	Sequence 11, Appl
30	475	8.1	4530	12	US-10-435-696-10	Sequence 208, App
31	475	8.1	4530	13	US-10-007-926A-119	Sequence 10, Appl
32	475	8.1	4530	13	US-10-101-510-124	Sequence 119, App
33	475	8.1	4530	13	US-10-338-730-1	Sequence 124, App
34	475	8.1	4530	13	US-10-116-275-131	Sequence 1, Appli
35	475	8.1	4530	13	US-10-272-437A-27	Sequence 131, App
36	475	8.1	4530	13	US-10-117-937-595	Sequence 27, Appl
37	475	8.1	4530	15	US-10-177-293-125	Sequence 595, App
38	475	8.1	9274	9	US-09-811-123-7	Sequence 125, App
39	475	8.1	9274	9	US-09-811-115-1	Sequence 7, Appli
40	474	8.1	3768	10	US-09-854-356-9	Sequence 1, Appli
41	474	8.1	3768	10	US-09-930-125-1	Sequence 9, Appli
42	474	8.1	3768	13	US-10-313-644-1	Sequence 1, Appli
43	474	8.1	4473	11	US-09-441-411-5	Sequence 1, Appli
44	474	8.1	4473	13	US-10-101-510-81	Sequence 5, Appli
45	474	8.1	4473	15	US-10-146-473-32	Sequence 81, Appl
						Sequence 32, Appl

ALIGNMENTS

RESULT_1
US-09-814-353-19356
; Sequence 19356, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF OVARIAN CANCER
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25

; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19356
; LENGTH: 5118
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 4, 5114, 5115, 5116, 5117, 5118
; OTHER INFORMATION: n = A,T,C or G
US-09-814-19356

Alignment Scores:
Pred. No.: 0 Length: 5118
Score: 4713.50 Matches: 912
Percent Similarity: 88.17% Conservative: 64
Best Local Similarity: 82.38% Mismatches: 122
Query Match: 80.44% Indels: 10
DB: 13 Gaps: 5

US-09-397-967A-16 (1-1099) x US-09-814-353-19356 (1-5118)

QY 1 MetAlaProSerGluGluThrProLeuLeuProGlnArgSerCysSerLeuSerSer 20
DB 96 ATGGACCTCCCAAGTGAAGAGAGCGCCCTGATCCCTCAGCGTTTCATGAGCTCTGTGCC 155
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlnArg 40
DB 156 ACGGAGCTGGTCCCTCGATGCTGCTGCCGCTCGGGCGCCGCCGCCAGCGC 215
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaAla 60
DB 216 CTATCTTCTCTTGGGACCACTTGGCTGAGGACCTGTGGTGAGCGTGCAGCGTCCAAAGGCC 275
QY 61 CysGlyLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
DB 276 ACGGAGCTCTGCTGTATCCACTCCCTCTTGTCTTGGCCACGAGGACCTGTCTGC 335
QY 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
DB 336 TGGTTCCTCCCGAGCACATCTTCCTCGTGAGGATGCCAGCACCCCAAGTCTGTGTAC 395
QY 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120
DB 396 AGGATTTCGCTTTTACTTCCCAATTGGTTTGGGCTGGAGAAGTGCCACCGCTTCGGGCTA 455
QY 121 ArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
DB 456 CGCAAGATTGGCCAGTGCTATCTTGTACCTGCCAGTCTCTGGAGCACCTCTTTGCCAG 515
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetIysGluGlnGly 160
DB 516 CACCGCAGTGACCTGGTGAGTGGCGCTCCCGCTGGGCTCAGTCTCAAGGAGCAGGT 575
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
DB 576 GAGTGTCTACGCTGGCGCTGTGGACCTGGCCGGATGGCGGAGAGCGGCCCGG 635
QY 181 ProGlyGluLeuLeuLeuThrValSerTyrLeuAlaCysLeuProProSerLeuArgAsp 200
DB 636 CCGGGAGAGCTGTGAAGATGTTCAGTACAGGCTGCTACCCCAAGCTTGGCGGAC 695
QY 201 ValIleGlnGlyGlnAspPheValThrArgArgIleArgThrValValLeuAla 220
DB 696 CTGATCCAGGCGCTGAGCTTCGTACCGGAGGCGTATTCGGAGGACGCTGCCGAGAGCC 755
QY 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaIysTyrIleLe 240
DB 756 CTGGC-CCGGCTGGCCCGCTGCCAGGAGACCGGACCTCGCTCATGCCCAAGTACATCAT 814
QY 240 uAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGln 260

DB 815 GGACCTGGAGCGGTGGATCCAGCGGGCGCGCCAGACCTTCCACGTGGCGCTCCCTGG 874
QY 260 yAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePr 277
DB 875 GGCCCTTGGTGGCCACAGCAGCGGCTGGGCTGCTCCGCGTGGCTGTGTGACGGCGGATCGC 934
QY 277 oTrpSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluLeuValAs 296
DB 935 CTGGACCCAGGAGAACAGGAGTCTCCAGCCCTTCTCGGACTTTCAGAAATCGTAGA 994
QY 296 pValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrVa 316
DB 995 CATTAGCATCAAGCAGGCGCCCGCGGCTGGCCCGCGGAGAGACCGGCTCGTCACTGT 1054
QY 316 lThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSe 336
DB 1055 TACCAGGACAGAACACAGATTTTAGAGGCCGAGTTCCAGGGCTGCGCCAGGCTCTGTC 1114
QY 336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCy 356
DB 1115 GTTCGTGGGCTCTGTGGACGGCTACTTCGGGCTGACACCGGACTCCACGACTTCTCTG 1174
QY 356 sIysGluValAlaProProArgLeuLeuGluAlaAspValCysHisGlyProI 376
DB 1175 CAAGGAGGTGGCACCGCGGAGCTGTGGAGGAAGTGGCCGAGCAGTGCACGGCCCAT 1234
QY 376 eThrLeuAspPheAlaIleHisIleLeuLeuAlaGlySerLeuProGlyThrTyrI 396
DB 1235 CACTCTGGACTTTGGCATCAAGCTCAAGACTGAGGCTGCTGCTCTCTCTCTGTGTGTCAGAACCC 1354
QY 396 eLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrPr 416
DB 1295 TCTCCGCGCAGCCCGGAGCTTTGACAGCTTCTCTCTCTCTCTCTGTGTGTCAGAACCC 1354
QY 416 oLeuGlyProAspTyrLysGlyCysLeuLeuArgGlnAspProSerGlyAlaPheSerIe 436
DB 1355 CCTTGGTCTGTATTAAAGGCTGCTCATCCGGCGCAGCCCGCCACAGGAACTTCTCTTCT 1414
QY 436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAalaCysTrpAsnSe 456
DB 1415 GGTGTGCTCAGCGGACCCCGCAGCAGTCTTGGAGAGTCTCTGGCAACCTCTGTGGATGG 1474
QY 456 rGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLy 476
DB 1475 GGGGCTGCAGTAGATGGGTGGCAGTCCCTCACTTCTGTGTATCCCGAGACCCAA 1534
QY 476 sGluLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCy 496
DB 1535 AGAAAGTCCAACTGATGCTGTGTCAGAGAGTCCAGCCCGCCACCATCATCTCTTGGT 1594
QY 496 sSerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSe 515
DB 1595 TCAGCCCGCATCCCAATACAGCTGAGTCAGATGACATTTCAAGATCCCTCTGTACAG 1654
QY 515 rLeuGluTyrHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerAr 535
DB 1655 CCTGGAGTGCATGAGAACCTGGGCTGCTTCCACAGATTTACCGGGCTGTCTG 1714
QY 535 gArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuIysValMetAspSe 555
DB 1715 CCATGAGTGTGGATGGGAGGCGCCGAAAGACAGAGAGTCTCTGAAGTCTATGATGC 1774
QY 555 rArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575
DB 1775 CAAGACCAAGACTCATGAGTGCATCTCTGGAGAGCGAGCTGTGATGAGCAAGTGT 1834
QY 575 rTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGl 595
DB 1835 GTACCGGCTACTCGTGTCTGCCAGCGGTGTGTCATGGCTGGAGACAGCAGCATGGTGA 1894
QY 595 nGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgIysArgGlyHisLeuValSe 615

Db 1895 GGAATTTGTACACCTGGGGCCATAGACATGATCTCGGAAAAAGTGGCCACCTGGTGC 1954
Qy 615 rAlaserTyrLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAs 635
Db 1955 AGCCAGCTGGAGCTGCGAGGTGGTCAACACAGCTGACCTTACGCCCTCAACTATCTGGAGGA 2014
Qy 635 pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyG1 655
Db 2015 CAAGGCCCTGCCCATGCAATGCTCTCGCCCGAAGGTGCTCCTGGCTCGGGAGGGGC 2074
Qy 655 yAspGlyAsnProProPheIleLysLeuSerAspProLysValSerProThrValLeuSe 675
Db 2075 TGATGGGAGCCCGCCCTTCATCAAGAGTGAGTGACCTCGGGGTACGCCCGCTGTGTTAAG 2134
Qy 675 rLeuGluMetLeuThrAspArgIleProTyrValAlaProGluCysLeuGlnGluAlaG1 695
Db 2135 CCTGGAGATGCTACCGACAGATCCCTGGGTGGCCCCCGAGGTCTCCGGGAGGGCGCA 2194
Qy 695 nThrLeuCysLeuGluAlaAspLysTyrGlyPheGlyAlaThrThrTyrGluValPheG1 715
Db 2195 GACACTTAGCTTGAAGCTGACAAAGTGGGCTTCGGCGCCACAGCTCGGGAGGTGTTAG 2254
Qy 715 nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAs 735
Db 2255 TGGCGTCAACATGCCCATCACTGACCTGGATCTCTGAAGAAACTCCAAATTTTATGAGA 2314
Qy 735 pGlnGlyGlnLeuProAlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMe 755
Db 2315 CCGCAGCAGCTGCGGCCCCCAAGTGGACAGAGCTGGCCCTGCTGATTCAACAGTGCAT 2374
Qy 755 tAlaTyrAspProGlyArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775
Db 2375 GGCCTATGAGCGGTCCAGAGGCCCTCTCTCCGAGCGGTGCTTCTGACCTCAATAGCCT 2434
Qy 775 vIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 795
Db 2435 CATCTCTTCAGACTATGAGTCTCTCTCAGACCCACACCTGGTGCCTCCAGCCTCGTGA 2494
Qy 795 pGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluG1 815
Db 2495 TGGGCTGTGG---AATGGTCCAGCTCTATGCTGCCAGACCCACAGCTCTTCGAGGA 2551
Qy 815 uArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCy 835
Db 2552 GAGACACCTCAAGTACATCTCACAGCTGGGCAAGGCACTTTGGCAGCTGGAGCTGTG 2611
Qy 835 sArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHi 855
Db 2612 CCGCTATGACCCGCTAGGCGCAATACAGGTGCTGCTGGCCGTGAAACAGCTGCAGCA 2671
Qy 855 sSerValProAspGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHi 875
Db 2672 CAGCGGCCAGACGACGAGGAGGACTTTCAGCGGGAGATTTCAGATCTCAAGACACTGCA 2731
Qy 875 sSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArg-GlnSerLeuA 895
Db 2732 CAGTGATTTTCATGTCAGATATCGTGTGTGAGTATGCGCCCGCGCCGCGCAGCTGCG 2791
Qy 895 rGluValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgG 915
Db 2792 GGCTGGTGCATGGAGTACTCGCCAGCGGCTGCTTGGCGGACTTCTCTGCGCGCCAGCCGCG 2851
Qy 915 ly---LeuHisThrAspArgLeuLeuLeuPheAlaTyrPgnIleCysLysGlyMetGluT 934
Db 2852 CGGCGCTTCGATGCCAGCGGCTCTCTTCATTCCTCGCAGATCTGCAAGGGCATGGAGT 2911
Qy 934 yrLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluS 954
Db 2912 ACCTGGGCTCCCGCGCTGCGTGCACCGGACCTGGCGCCGCCGCAACATCTCTGTGGAGA 2971
Qy 954 erGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysA 974
Db 2972 GCGAGGGCACAGTCAGATCGCTGACTTCGGCCCTAGCTTAAGCTGCTGCCCGCTTGACAAG 3031

Qy 974 spTyrTyrValValArgGluProGlyGlnSerProIlePheThrTyrAlaProGluSerL 994
Db 3032 ACTACTACGTGGTCCGGAGCCAGGCGCAGAGCCCATTTTCTGGTATATGCCCCGGAATCCC 3091
Qy 994 euSerAspAsnIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrG 1014
Db 3092 TCTCGGACAAATCTTCTCTGCCAGTCAGAGCTCTGGAGCTTGGGGTCTCTCTGTACG 3151
Qy 1014 luLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetG 1034
Db 3152 AGCTCTTCACCTACTGCGACAAAAGCTGCAGCCCTCGGCCAGTTCCTCTCGGATGATGG 3211
Qy 1034 lyProGluArgGluGlyProProLeuCysArgLeuLeuLeuLeuAlaGluGlyArgA 1054
Db 3212 GATGTAGCGGAGTGTCCCCGCCCTCTGCCGCCCTCTTGGAACTCTGGAGAGGGCCAGA 3271
Qy 1054 rgLeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpA 1074
Db 3272 GGCTGCCGGCGCTCTCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3331
Qy 1074 laProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1094
Db 3332 CCCCTAGCCACAGACCGGCCATCAATTCAGCGCCCTGGGCCCCAGCTGGACATGCTGT 3391
Qy 1094 rPArgGlyArgProGly 1099
Db 3392 GGAGCGGAAGCCGGGG 3408

RESULT 2

US-10-305-720-1055
; Sequence 1055, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; PRIOR FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1055
; LENGTH: 3620
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank ID No. US20040010136A1 g1039418
US-10-305-720-1055

Alignment Scores:

Pred. No.:	0	Length:	3620
Score:	4597.00	Matches:	895
Percent Similarity:	86.80%	Conservative:	65
Best Local Similarity:	80.92%	Mismatches:	123
Query Match:	78.45%	Indels:	24
DB:	12	Gaps:	7

US-09-397-967A-16 (1-1099) x US-10-305-720-1055 (1-3620)

Qy 1 MetAlaProSerSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
Db 96 ATGGCACTTCAAGTGAAGAGAGCGCCCTGATCCCTCAGCGTTTCATGACGCTCTTGTGCC 155
Qy 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
Db 156 ACGAGAGCTGTGCTCCCTGATGTGCTGCCCGCTCGGGCCCGGGCCCCCGCCAGCGC 215
Qy 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
Db 216 CTATCTTCTCTTTGGGAGCCACTTGGCTGAGGACCTGTGTGCTGAGGCTGCCAAGGCC 275

QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
DB 276 AGCGGCAATCCCTGCTGTGTACACATCCCTCTTTGTCTGGCCACGGAGACCTGCTCTGC 335
QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
DB 336 TGGTTCCCGCCGAGCCACATCTTCTCCGTGGAGGATGCCAGACCCCAAGTCTCTGTGTAC 395
QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
DB 396 AGGATTCCGCTTTTACTTCCCAATGTTGGCTGGAGAGTGCCACCGCTTCGGGCTA 455
QY 121 ArgIysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
DB 456 CCGAAGGATTTGGCCAGTGCTATCTTGACCTGGCCAGTCTCTGGAGACCTCTTTGCCAG 515
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetIysGluGlnGly 160
DB 516 CACCGCAGTGACCTGGTGGTGGCGCCCTCCCGTGGGCTCAGTCTCAAGGAGCAGGCT 575
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaArg 180
DB 576 GAGTGTCTCAGCTCGCCCTGTGGACCTGGCCCGGATGGCGGAGAGAGGCCAGCGG 635
QY 181 ProGlyGluLeuLeuIysThrValSerTyrIysAlaCysLeuProProSerLeuArgAsp 200
DB 636 CCGGAGAGCTGCTGAAGACTGTACAGTACAAAGGCTGCTTACCCCAAGCTTGGCCGAC 695
QY 201 ValIleGlnGlyGluAsnPheValThrArgArgArgIleArgArgThrValValLeuAla 220
DB 696 CTGATCCAGGGCTGAGCTTCTGAGCGCGAGCGGTATTCGGAGGACGCTGGCGAGGCC 755
QY 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaIysTyrIleLeu 240
DB 756 CTGGC-CCGCGTGGCGCTGTCAGCGACAGCCGGCTCTGCTCATGGCCAAAGTACATCAT 814
QY 240 uAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260
DB 815 GGACCTGGAGCGCTGGATCCAGCGGGGCGCGGAGACCTTCCAGCTGGCGCTCCCTGG 874
QY 260 yAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePr 277
DB 875 GGCCTTGTGTGGCCACGACGGCTGGGCTGCTCCGCTGGTGGTGACGGCGCATCGC 934
QY 277 oTrpSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAs 296
DB 935 CTGGACCCAGGGAGAACAGGAGTCTCCAGCCCTTCTGCGACTTTCCAGAAATCGTAGA 994
QY 296 pValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrVa 316
DB 995 CARTAGCATCAAGAGGCCCCGCGCTGGCCCGCGGAGACACCGCTGGTCTACTGT 1054
QY 316 lThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSe 336
DB 1055 TACCAGGACAGACACACAGATTTTAGGCGCGAGTTCCCGAGGCTGCGCGAGCTCTGTC 1114
QY 336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCy 356
DB 1115 GTTCGTGGCGCTCGTGGAGCGCTACTTCCGCGTGACCCAGGACTCCCGAGCACTTCTCTG 1174
QY 356 sIysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIl 376
DB 1175 CAAAGAGGTGGCACCGCGAGGCTGTGGAGGAAGTGGCCGAGAGTGGCCACCGGCCCAT 1234
QY 376 eThrLeuAspPheAlaIleHisIysLeuLysAlaAlaGlySerLeuProGlyThrTyrIl 396
DB 1235 CACTCTGGACTTTGCCATCAACAAGCTCAAGACTGGGGGCTCAGCTCTGGCTCTATGT 1294
QY 396 eLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrPr 416
DB 1295 TCTCCGCGGACCGCCAGGACTTTGACGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1354

QY 416 oLeuGlyProAspTyrIysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerIle 436
DB 1355 CTTTGGTCTCTGATTATAAGGCTGCTCATCGCGCGCAGCCCCACAGGAACCTCTCTTCT 1414
QY 436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSe 456
DB 1415 GGTGGCTCAGCCAGCCACAGCAGTCTTCGAGAGCTCCCTGGCAACCTCTCTGGGATGG 1474
QY 456 rGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProly 476
DB 1475 GGGGTGCAGTAGATGGGTGGGAGTCCCTCATCTCTCTCTCTCTCTCTCTCTCTCTCT 1534
QY 476 sGlyIysSerAsnLeuIleValValArgGlyCysAsnProAlaProAlaProGlyCy 496
DB 1535 AGAAAGTCCAACTGATCGTGTCCAGAGAGTCCAGCCCCACCCATCATCTCTTGT 1594
QY 496 sSerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAsnSe 515
DB 1595 TCAGCCCAATCCCAATACAGCTGAGTCAGATGACATTTCAAGATCCCTGTGTGACAG 1654
QY 515 rLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerAr 535
DB 1655 CTGGAGTGGCATGAGAACCTGGGCTCTTCCACCAAGATTTACCGGGGCTGTCTG 1714
QY 535 gArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSe 555
DB 1715 CCATGAGTGTGGATGGGAGGCCCGAAAGACAGAGTGTCTGTAAGTCTATGATGC 1774
QY 555 rArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575
DB 1775 CAAGCACAGAACTGCATGGAGTCAATCTCTGGAACAGCGAGCTTGATGAGCAAGTGC 1834
QY 575 rTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGl 595
DB 1835 GTACCGGCTATCTGCTGCTCCAGCGCTGTGCATGGTGGAGACAGCACCATGGTGA 1894
QY 595 nGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSe 615
DB 1895 GGAAATTTGTACACCTGGGGGCCATAGACATATATCTCGGAAACACGTGGGCACCTGTGTGC 1954
QY 615 rAlaSerTrpLysLeuGlnValThrIysGlnIleAlaTyrAlaLeuAsnTyrLeuGluAs 635
DB 1955 AGCCAGCTGGAGCTGACAGTGGTCAACAGCTGGCTACGCCCTCACTATCTGGAGGA 2014
QY 635 pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGl 655
DB 2015 CAAAGGCTTCGCCCATGGCAATGTCTCTGCCGGAAGTGTCTCTGGCTCGGAGGGGC 2074
QY 655 yAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSe 675
DB 2075 TGATGGAGCCCGCCCTTTCATCAAGCTGAGTGACCTGGGGTTCAGCCCGCTGTGTTAG 2134
QY 675 rLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGl 695
DB 2135 CTTGGAGATGCTCACCGACAGATCCCTGGGGTGGCCCGAGTGTCTCGGAGGGCGCA 2194
QY 695 nThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGl 715
DB 2195 GACACTTAGCTTGGAGAGCTGACAACTGGGCTTCGGCGCCACCGTCTGGGAAGTGTAG 2254
QY 715 nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAs 735
DB 2255 TGCGCTCACCATGCCCTCAGTCCCTGGATCTCTGCTAAGAACTCAATTTTATGAGGA 2314
QY 735 pGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMe 755
DB 2315 CCGGACAGAGCTCCCGGCCCGCCCAAGTGGAGACAGAGCTGGCCCTCTGATTCACAGTGCAT 2374
QY 755 tAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775
DB 2375 GGCCTATGAGCGGTCCAGAGGCCCTCTCTCCGAGCGCTCATTCGTGACCTCAATAGCT 2434
QY 775 uIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 795

[illegible]

RESULT 3

US-10-191-803-200 ; Sequence 200, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna


```
Db 2887 GTGAAGTCAAGGGAGTGTGTACAGTGTGCTGGCGCAACCTAAGATTAATTATGGAA 2946
Qy 900 TyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHisThr 918
Db 2947 TATTTGGCGTATGGAAGTTTAGGAGACTATCTCCAAAACATATAAGAACGGATAGATCAC 3006
Qy 919 AspArgLeuLeuLeuPheAlaTrpGlnLeCysLysGlyMetGluTyrLeuGlyAlaArg 938
Db 3007 AAAAAAAGTCTTCGATATACATATCCAGATATGCAAGGCGATGAGTATCTTGGTACAAA 3066
Qy 939 ArgCysValHisArgAspLeuAlaAlaArgAsnLeuLeuValGluSerGluAlaHisVal 958
Db 3067 AGGTATATCCACCGGAGTCTGCAACAAGGAATATATTGGTAGAAAATGAGAACAGAGTT 3126
Qy 959 LysileAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVal 978
Db 3127 AAAATAGGAGATTTTGGATTAAACCAAGTCTTGCCACAGGCAAGAATACTTACAAAGTA 3186
Qy 979 ArgGluProGlyGlnSerProilePheTrpTyrAlaProGluSerLeuSerAspAsnIle 998
Db 3187 AAGGAACACAGGTGAAGTCCCATATTTCTGGTATGCTCCGAATCTTGACAGAGCAAG 3246
Qy 999 PheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
Db 3247 TTTTCTGTGGCTCAGATGTGGAGCTTTGGTGGTCTTGTACGAACTTTTTCACGTAT 3306
Qy 1019 CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGlu 1038
Db 3307 ATCGAGAAGACAAAGCCACCTGTGGAATTTATGCGAATGATCGGCAATGATNAACAA 3366
Qy 1039 GlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProPro 1057
Db 3367 GGGCAGATGATCGTATTCCATTTGATAGAGCTACTGAAGAACAAACGGAAGATCGCGAGA 3426
Qy 1058 ProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGluPro 1077
Db 3427 CCAGAGGTCCTCCAGACGAGATTTATGTGATCATGACAGAAATGCTGGAACAACAAATGC 3486
Qy 1078 HisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093
Db 3487 AACCAAGTCTCTTTTCAGGAGCTTTCCCTTCGGGTGGATCAATC 3534

RESULT 4
US-10-305-720-297
; Sequence 297, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 297
; LENGTH: 4078
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040010136A1 179527
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (4078)
; OTHER INFORMATION: a, t, c, g, or other
US-10-305-720-297

Alignment Scores:
Pred. No.: 3,83e-257 Length: 4078
Score: 2594.00 Matches: 531
Percent Similarity: 65.94% Conservative: 193
```

```
Best Local Similarity: 48.36% Mismatches: 343
Query Match: 44.27% Indels: 32
DB: 12 Gaps: 14
US-09-397-967A-16 (1-1099) x US-10-305-720-297 (1-4078)

Qy 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ArgGlyProGlyProGlyProGlyProGly 39
Db 447 TCCAGTTCTTCAGGT-----GTATCTTTACCATTTCCCTTGGGAAATCTGAGGCAGA 497
Qy 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLys 59
Db 498 TTATCTGACCTTTCCATCTGGGAGTATGTTGGAGAAGAATCTGATTGCTGCTCTTAA 557
Qy 59 sAlaCysGlyLysLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSe 79
Db 558 AGCTTGTGTATCATCACCTGTGTATCATTAATATGTTGCTTTAATGAGTAAGACAGAAAG 617
Qy 79 rCysTrpPheProProSerHisIlePheCysTrpLeuGluAspValAspThrGlnValLeuVa 99
Db 618 GATCTGGTATCCACCAACCAATGCTTCCATATAGATGAGTCAACCAAGGCATAATGTACT 677
Qy 99 lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr 117
Db 678 CTACAGAATAAGATTTTACTTTCTCTCG-TGGTATTGCAATGCGACCAACAGAGCCTATCG 736
Qy 117 gPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137
Db 737 GCATGGAAATATCTCGAGGTGCTGAAGCTCTCTCTTCTGATGACTTGTGTCATGCTTACCT 796
Qy 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157
Db 797 CTTTGTCTCAGTGGCGGCGATGATTTTGTGATGGATGGAATAAAGTACCTGTGACTCATGA 856
Qy 157 sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl 177
Db 857 AACACAGGAAGAATGCTTCTGGATGACAGTGTGATGATGATGATGATGATGATGATGATGAT 916
Qy 177 nAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSe 197
Db 917 CGATCAAAACCCCACTGGCCATCTATACTCTATCAGCTACAAGACATTTCTTACCACAATG 976
Qy 197 rLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgThrVa 217
Db 977 TATTGAGCAAAAGATCCAGACTATCATATTTTTCACAAAGGAAGCAATAAAGGTACAGATT 1036
Qy 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLy 237
Db 1037 TCGCAGATTTTATTCAGCAATTCAGCCCAATGCAAGCCCACTGCCAGAACTTGAAACTTAA 1096
Qy 237 sTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGl 257
Db 1097 GTATCTTATAAATCTGGAAACTCTGCAGTCTGCTCTTACACAGAGAAATTTGAAGTAA 1156
Qy 257 yLeuProGlyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAs 273
Db 1157 AGAACCTGGGAAGTGGCTCTTCAGGTGAGGAGATTTTTCGAACCATTAATAACTGAAA 1216
Qy 273 pAsnGlyIleProTrpSerSer-----AsnAspGluLe 284
Db 1217 CGGTGGAATTCAGTGTCAAGAGGGAACATAAAGAAAGTGACACACTGACAGAACAGGA 1276
Qy 284 uPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProAr 304
Db 1277 TTTACAGTTTATATGCGATTTTCTTAATAATTATTGATGTGATGATTAAGCAAGCAACCA 1336
Qy 304 gValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLe 324
Db 1337 AGAGGCT---TCAAATGAAGCCGAGTTGTAATCTATCCATAAGCAAGATGGTAAAAATCT 1393
Qy 324 uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTy 344
Db 1394 GGAATTAAGTACTAGCTCATTAAAGGAAGCTTTGTCTTCTGTCGTCAATTAATGATGATA 1453
```


Db 2087 ATATGAACCAATGGTGTTCACAAAATCAGAAATGGAATTTGATATTTAATGAAGCC 2146
 Qy 522 euGlyHisGlySerPheThrIysLlePheArgGlySerArgGluValValAlaSp---G 541
 Db 2147 TTGGCCAGGCACTTTTACAAAGATTTTAAAGGCGCTACGAAGAGAGTAGGAGACTAGC 2206
 Qy 541 lYgluThrHisAspSerGluValLeuLeuValMetAspSerArgHisArg-AsnCys 560
 Db 2207 GTCAATGTCATGAACAGAGTTCTTTTAAAGTTCTGGATTAAGCAGACAGAAACTAT 2266
 Qy 561 MetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuVal 580
 Db 2267 TCAGAGTCTTTCTTGAAGCATCAAGTAGATGAGCAAGCTTCTCACAAGCATTTGGTT 2326
 Qy 581 LeuLeuHisGlyValCysMetAlaGlyAsp---SerIleMetValGlnGluPheValTyr 599
 Db 2327 TTAATAATGAGTAGTGTCTGTGAGAGCAGAAATATTCTGGTTCAGGAGTTTGTAAA 2386
 Qy 600 LeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLys 619
 Db 2387 TTTGGATCACTAGATACATATCTGAAAGAAATPAAATTTGTATAATATATATGAA 2446
 Qy 620 LeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuPro 639
 Db 2447 CTTGAAGTTGCTTAAACAGTTGGCAGCGGCATGCAATTTCTAGAGAAACACCCCTATT 2506
 Qy 640 HisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGly-----AspGly 657
 Db 2507 CATGGCAATGTATGTCACAAACATATTCTGCTTATCAGAGAGAGAGAGAGAGAGAG 2566
 Qy 658 AsnProPheIleLysLeuSerAspProGlyValSerProThrValLeuLeuLeuGlu 677
 Db 2567 AATCCCTCTTCACTAACTTAGTAGTCCTGGCATTTAGTATTAAGTTTGGCAAGGAC 2626
 Qy 678 MetLeuThrAspArgIleProTyrValAlaProGluCysLeuGlnGluAlaGlnThrLeu 697
 Db 2627 ATTCTTCAGAGAGAGATACCATGGGTACCACTGATGATGATGAAATCCTAAATTTA 2686
 Qy 698 CysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGly 717
 Db 2687 AATTGGCAACAGACAAATGAGTTTGGTACCACCTTTGTGGAAATCTGCAGTGGAGGA 2746
 Qy 718 Pro-AlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGly 737
 Db 2747 GATAATACCTCTAGTGTCTGATCCCAAGAGAGCTACAAATTTATGAGATAGGCA 2806
 Qy 737 yGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTy 757
 Db 2807 TCAGCTTCTGCACCAAGTGGGAGAGATTACACACCTTATAAATAATTTGATGGATTA 2866
 Qy 757 rAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleTh 777
 Db 2867 TGAACCAAGATTTCAGGCGCTCTTTCAGAGCCATCATACAGATCTTAACAGTTTGTATTAC 2926
 Qy 777 rSerAspTyrGluLeuLeuSerAspProThrProGlyLleProSerProArg---AspGl 796
 Db 2927 TCCAGATTATGAATTAACAGAA---AATGACATGTTACCAATATGAGATAGGTGC 2983
 Qy 796 uLeuCysValAlaAlaGlyAlaGlnLeuTyrAlaCysGln-AspProAlaIlePheGluGlu 816
 Db 2984 CCTAGGGTTTCTGGTACC-----TTTGAAGACCGGGATCTCTACAGTTTGAAGAGA 3037
 Qy 816 rgHisLeuLysTyrIleSerLeuLeuLysGlyAsnPheGlySerValGluLeuCys- 835
 Db 3038 GACATTGAAATTTCTACAGCAACTTGGCAAGGGTAATTTTGGGAGTGTGGAGATGTGCC 3097
 Qy 836 ArgTyrAspProLeuGlyAsnThrGlyProLeuValAlaValLysGlnLeuGlnHis 855
 Db 3098 CGGTATGACCTCTACAGCAACACTGGGAGGTGCTGCTGTAAAGAAAGCTTCAGCAT 3157
 Qy 856 SerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHis 875
 Db 3158 AGTACTGAAGAGCACCTTAAGAGACTTTGAAGGGAAATTTGAATCTCGAAATCCCTACAG 3217

Qy 876 SerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArg 895
 Db 3218 CATGACAACTTGTAAAGTACAAAGGAGTGTCTACAGTGTGTGGTGGCGTAATCTAA 3277
 Qy 896 LeuValMetGluTyrLeuPro-SerGlyCysLeuArgAspLeuLeuGlnArgHisArg-- 914
 Db 3278 TTAATATGGAATATTTTACCATATGGAAGTTTACGAGACTATCTTCAAAAACATAAGA 3337
 Qy 915 -GlyLeuHisThrAspArgLeuLeuPheAla-TripGlnIleCysLysGlyMetGluT 934
 Db 3338 ACGGATAGATCATATAAACTTCTGCAGTACACATCTCCAGATATGCAAGGGTATGGAGT 3397
 Qy 934 yTrpGlyAlaArgCysValHisArgAspLeu-AlaAlaArgAsnIleLeuValGlu 953
 Db 3398 ATCTTGTACAAAAGGTATATCCACAGGATCTGGCAACGAGAAATATTTGGTGGAG 3457
 Qy 954 SerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuProLeuGlyLys 973
 Db 3458 AACGAGAACAGAGTTTAAATTTGGGTTAACCAAGTCTTGCACAGACAA 3517
 Qy 974 AspTyrTyrValValArgGluProGlyGlnSerProIlePheTyrTrpTyrAlaProGluSer 993
 Db 3518 GAATCTATAAGTAAAGAACCTGTGTGAAGTCCCATATTTCTGGTATGCTCCAGATCA 3577
 Qy 994 LeuSer-AspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTy 1013
 Db 3578 CTGAGCNGAGAGCAAGTTTCTGTGGCTCAGATGTTTGGAGCTTTGGAGTGTCTGTGA 3637
 Qy 1013 rGluLeuPheThrTyrCysAsp-LysSerCysSerProSerAlaGluPheLeuArgMet- 1032
 Db 3638 TGAACCTTTTACATACATTTGAGAAAGTAAAGTCCACAGCGGAATTTATCGGTATGA 3697
 Qy 1033 MetGlyProGluArgGluGlyProProLeuCys---ArgLeuLeuGluLeuAlaGlu 1051
 Db 3698 TTGGGCAATGACAAACAGAGCAGAGATGATCGTCTTCCATTTGTAGAACCTTTTGAAGAAT 3757
 Qy 1052 GlyArgGluProProProProThrCysProThrGluValGlnGlu-LeuMetGlnLe 1071
 Db 3758 AATGGAAGATTACCAAGACAGATGGATGCCAGATGAGATGATGATGATGATGATGATG 3817
 Qy 1071 uCysTrpAlaProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAs 1091
 Db 3818 ATGCTGAAACATAATGTAATCAACGCCCTCTTTAGGATCTAGCTCTTCGAGTGA 3877
 Qy 1091 pProLeu 1093
 Db 3878 TCAATA 3884

RESULT 6

US-10-295-027-697
 ; Sequence 697, Application US/10295027
 ; Publication No. US2003023250A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Afar, Daniel
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsberg, Wendy M.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Hevez, Peter A.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Watson, Susan R.
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
 ; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
 ; FILE REFERENCE: 018501-012500US
 ; CURRENT APPLICATION NUMBER: US/10/295,027
 ; CURRENT FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: US 09/663,733
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13

QY 483 ValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysCysAla 502
Db 1941 GTGGCTACTAAG-----AAAGCCAGAGGTGGAGCC--GTTACCCC 1982
QY 503 LeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeu 522
Db 1983 ATGAGCCAGCTGAGTTTCGATCGATCTCTCAAGAGGATCTGGTCGAGGGGAGACACCTT 2042
QY 523 GlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValValAspGlyGlu 542
Db 2043 GGGAGAGCAGAGAACACACATCTATTCTGGGACCTGTATGATTAACAGGATGACGAA 2102
QY 543 ThrHisAspSerGlu-----ValLeuLeuLysValMetAspSerArgHisArg 558
Db 2103 GGAATCTTCTGAAGAGAGAGAGATAAAGTGTATCTCAAGTCTTAGACCCAGCCACAGG 2162
QY 559 AsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTrpProHis 578
Db 2163 GATATTTCCCTGGCTTCTTCGAGGAGCAGCATGATGAGACAGGTCTCCACACAAACAC 2222
QY 579 LeuValLeuLeuHisGlyValCysMetAlaGly----AspSerIleMetValGlnGluPhe 597
Db 2223 ATCGTGTACCTCTATGGGTCTGTGTCTGGCAGCGTGGAGATATCATGTGGAGAGTTT 2282
QY 598 ValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSer 617
Db 2283 GTGGAGGGGGTCTCTGATCTCTTCATGTCACCGGAAAGCGATGCTTACACACCA 2342
QY 618 TrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGly 637
Db 2343 TGGAAATTCAAAGTTGCCAAACAGCTGGCCAGTGCCTGAGCTACTTGGAGGATAAAGAC 2402
QY 638 LeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGlyAspGly 657
Db 2403 CTGTGCTCAATGGAATGTGTACTAAACCTCTCTGGCCGCGAGGGGCATCGACAGT 2462
QY 658 Asn---ProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeu 676
Db 2463 GAGTGTGGCCATTCATCAAGCTCAGTACCCCGGATCCCATACGCTGTGCTAGG 2522
QY 677 GluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThr 696
Db 2523 CAAGATGATGATGAACGATCCATGATGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 2582
QY 697 LeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArg 716
Db 2583 CTGAGTGTGGCTGCTGACAAAGTGGAGCTTTGGAAACCCAGCTCTGGGAAATCTGCTACAAT 2642
QY 717 GlyProAlaHisIleThrSerLeuGluProAlaLysLeuLysPheTyrGluAspGln 736
Db 2643 GCGAGATCCCTTTGAAAGACAGACGCTGATGGAAGAGAGATCTTATGAAAGCCGG 2702
QY 737 GlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAla 756
Db 2703 TGCAGCCAGTGCACACCATCATGATGAGAGCTGGCTGACCTCATGACCCGCTGCATGAAC 2762
QY 757 TyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIle 776
Db 2763 TATGACCCCAATCAGAGGCGCTTCTTCGAGCCCATCATGAGACATTAATAAGCTTGAA 2822
QY 777 ThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGlu 796
Db 2823 GACAGATCCAGATATGTTTTCAGAAAAAACCCAGCACTGAA----- 2867
QY 797 LeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArg 816
Db 2868 -----GTGACCCCAACACATTTTGAAGAGCGC 2894
QY 817 HisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArg 836
Db 2895 TTCTTAAGAGGATCCGTGATCTTGGAGAGGGCCACTTTTGGGAAGGTTGAGTCTGCGAG 2954

QY 837 TyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGln---His 855
Db 2955 TATGACCCCAAGGGGCAATAACAGGGGAGCAGGTGCTGTAAATCTCTGAAGCCCTGAG 3014
QY 856 SerValProAspGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHis 875
Db 3015 AGTGGAGTACCAACATAGCTGATCTGAAAAGGAATTCGAGATCTTAAGAACTCTAT 3074
QY 876 SerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArg 895
Db 3075 CATGAGAACATTTGTGAAGTACAAAGGAATCTGCACAGAACGAGGAAATGGTATTAAAG 3134
QY 896 LeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly 915
Db 3135 CTATCATGGAATTTCTGCTTCGGGAAGCTTAAAGGAATATCTCCAAAGAAATGAAGAAC 3194
QY 916 ---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyr 934
Db 3195 AAAATAAACCTCAACAGCAGCTAAATATATGCCGTTGAGATTTGTAAGGGGATGAGCTAT 3254
QY 935 LeuGlyAlaAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSer 954
Db 3255 TTGGGTTCTCGCAATACGTTTACCGGAGCTTGGCAGCAAGAAATGTCTCTTGTGAGAGT 3314
QY 955 GluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAsp 974
Db 3315 GAACACCAAGTGAATAATTTGGAGACTTCGGTTTAAACCAAGCAATGAACCCGATGAAGAG 3374
QY 975 TyrTyrValValArgGluProGlyGlnSerProIlePheTyrTrpAlaProGluSerLeu 994
Db 3375 TATTACACCGTCAGGATGACCGGACAGCCTGTGTTTGGTATGCTCCAGATGTTTA 3434
QY 995 SerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGlu 1014
Db 3435 ATGCAATCTAAATTTTATATTGCTCTGACGCTGCTCTTTTGGAGTCACTCTGCATGAG 3494
QY 1015 LeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034
Db 3495 CTGTGACTTACTGTGATTCAGATTCTAGTCCCATGGCTTTGTTCTCTGAAATGATAGC 3554
QY 1035 ProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgArg 1054
Db 3555 CCAACCCATGGCAGATGACATGACATGACATGTAATACCTTAAAGAGAGAAACGC 3614
QY 1055 LeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAla 1074
Db 3615 CTGCGGTGCGCCACCTAACTGTCCAGATGAGTTTATCAACTTATGAGGAAATGCTGGAA 3674
QY 1075 ProGluProHisAspArgProAlaPheAlaThrLeu 1086
Db 3675 TTCCAACCATCCAATCGGACAAAGCTTTTCAGAACCTT 3710

RESULT 7

US-10-084-817-230
; Sequence 230, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2002-02-25
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 230
; LENGTH: 5073
; TYPE: DNA
; ORGANISM: Homo sapiens

QY 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
DB 528 TGTACGCTCCGACCGAATCATCACTGTGGATGACAAAACGCTCTCTCGGCTCCACTAC 587
QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeu-----GluThrCysHis 116
DB 588 CGCATGAGGTCTACTTTACCAACTGGCAGCGAACCAATGACAAACAGCTCTGTATGG 647
QY 117 ArgPheGlyLeuValGlyAspLeu-----ThrSer 126
DB 648 CGACATTCCTCAAGAGACGAGAAAACCGCTATGAGAGAAAGGTTCCGAGAACCAACC 707
QY 127 AlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHisArgSerAspLeuVal 146
DB 708 CCATCTCTGTATGTCAGTTCACTGGAGTATCTGTTGACAGGACAGTATGTTGATC 767
QY 147 SerGlyArgLeuProValGlyLeuSerMetIysGluGln-----Gly 160
DB 768 AAATGCTGCTCCATTTCCGGACCCCAAGACGGAGCAAGACGGACATGATATTGAAAT 827
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
DB 828 GAGTGCCTGGCATGGGGTCTCTGGCCATCTCCCACTATGCCATGATGATGAGAAAGATGCAG 887
QY 181 ProGlyGluLeuLeuLysThrValSerTyrIysAlaCysLeuProProSerLeuArgAsp 200
DB 888 TTGCGCGAACTTCCCAAGACATCAGCTACAGCGATATATTCCAGAAACATTTGAATAA 947
QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgIle----- 213
DB 948 TCATCAGACAGAGAACTTCTTACAGGATGCGAATAATAATGTTTCAAGGATTTTC 1007
QY 214 -----ArgArgThrValValLeuAlaLeuLeuProCysGlyArgLeuPro 228
DB 1008 TTGAAGGAATTTAAACAAGACCATC-----TGTGAC 1040
QY 229 GlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGlu-----ArgLeuHis 246
DB 1041 AGCAGTGTGCATGACCTGAAGTGAATACCTGGCTACCTTGGAACTTCTACATTGACA 1100
QY 247 ProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGluProGly 266
DB 1101 AACATTTATGAGCTGAATATTGTAGACTTCTATGCTACTGATTTTCATCAGAAATGAA 1160
QY 267 LeuLeuArg-----ValAlaGly 272
DB 1161 TTGAGTCGATGCTCATGATGACGTGGCAATGTTCTCTATGAGTCTGCTGGTGGGA 1220
QY 273 AspAsnGlyIleProTrpSerSerAsnAsp----- 282
DB 1221 AATCTCGGATCCAGTGGCGGAGAAACCAAAATGTTGTTCTCTGTGAAAGGAAATAAT 1280
QY 282 ----- 282
DB 1281 AAACCTAAGCGGAAAACCTGGAATATAATAACACAAAGAGATGATGAGAAACAAA 1340
QY 283 -----GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsn 300
DB 1341 CTCCGGGAAGAGTGGAACTTTTCTTATTTCCCTTGAATATCCCAACATTTGTAATAAG 1400
QY 301 GlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAsp 320
DB 1401 GAGTCT-----GTGTCAGCATTAACAACAGGAC 1430
QY 321 GlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeu 340
DB 1431 AACAAAAACATGGAATCTCAAGTCTCTCTTCGAGAGGAAGCCTTGTCTTGTGCTCCCTG 1490
QY 341 ValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysGlyGluValAla 360
DB 1491 GTGGATGGTACTTCCGGCTCACTGCAGATGCCACCATTTACTCTGTACTGTGGCT 1550
QY 361 ProProArgLeuLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPhe 380

DB 1551 CCCCACCTGATGTCCACAATATACAGACGGCTCCACGGTCCCAATCTGCACAGATAT 1610
QY 381 AlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSer 400
DB 1611 GCCATCAATAAGCTGCGCAGGAGGAGTGAAGAGGGATGTACGTCTGAGGTGGAGC 1670
QY 401 ProGlnAspTyrAspSerPheLeuLeuThr-----AlaCysValGlnThrPro-----Leu 417
DB 1671 TGCACCGACTTTGACACATTTCTATGACTGTCACTGCTTGAAGAGTCTGAGGTATTG 1730
QY 418 Gly-----ProAspTyrLysGlyCysLeuLeuArgGlnAspProSerGlyAlaPheSer 435
DB 1731 GGTGGCCAGAACGAGTTCAAGAACTTTTCAGATT---GAGGTACAGAGGGCGCTTACAGC 1787
QY 436 LeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsn 455
DB 1788 CTGCATGGCTCTATGGACCACTTTCCAGCCTCGGAGACCTCATGAACCACTCTAGAG 1847
QY 456 SerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgPro 475
DB 1848 CAGATCTCTGCGCAGCAGACACATAAGCTTTGTGCTGAAACGATGCTGTCAAGCTT 1907
QY 476 LysGluLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGly 495
DB 1908 CGAGAAATCTCCAATCTGCTGTCAGCCATAAG-----AAAGCCCAAGAG 1952
QY 496 CysSerProSerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSer 515
DB 1953 TGGCAGCCT---GTCATCTCATGAGCAGCTGAGCTTTGATCGGATCTTTAAGAAAGAT 2009
QY 516 LeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArg 535
DB 2010 ATTATACAGGTGAGCAGCTTTGCGAGGACACAAAGACACATATCTATTCTGGACCTG 2069
QY 536 ArgGluValValAsp-----GlyGluThrHisAspSerGluValLeuLeuLys 551
DB 2070 CTGAGTACAGAGTGAAGAGAAATTTGCTGAAGAGAGAAAGATATAAGATGATCTCAA 2129
QY 552 ValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMet 571
DB 2130 GTCTAGACCCAGCCAGCCGGACATCTCTCGGCTTTCTTTGAGGCTGTGATGATG 2189
QY 572 SerGlnValSerTyrProHisLeuValLeuHisGlyValCysMetAlaGly---Asp 590
DB 2190 AGACAGGTTTCCCAACACATATAGTACTCTACGGCTGTGTGTCGAGATGTGAA 2249
QY 591 SerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArg 610
DB 2250 AATATCATGTGGAGAGATTTGTGGAGGGGGCGCTTGGATCTCTTCATGCACCGGAAA 2309
QY 611 GlyHisLeuValSerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeu 630
DB 2310 AGTGATGCGCTTACTACCCCTGGAGTTCAAGTTGCCAAACAGCTGGCCAGTGGCCTG 2369
QY 631 AsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeu 650
DB 2370 AGTTACTTGAAGATAAAGACCTGTTTCATGGAATGTGTGCACATAAAACCTCTCTG 2429
QY 651 AlaArgGluGlyGlyAspGlyAsn---ProProPheIleLysLeuSerAspProGlyVal 669
DB 2430 GCCCTGAGGGATTTGACAGTGCATTTGCCCGCTTCATCAAGCTTAGTACCTGTCATC 2489
QY 670 SerProThrValLeuSerLeuGluMetLeuThrAspArgIleProTrpValAlaProGlu 689
DB 2490 CCAGTCTCTGTCTCACCAGGCAAGTGCATAGAGCAATCCCTGGATGCTCTGTAG 2549
QY 690 CysLeuGlnGluAlaGlnThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThr 709
DB 2550 TGTGTTGAAGACTCCAGAACTGAGTGTGGCTGTGTGACAGTGGAGCTTTGGAAACAGC 2609
QY 710 ThrTrpGluValPheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaLysLys 729

Db 2610 CTCTGGGAATCTCTCTACACGGAGAGATTCTCTCTCAAGACAGACCCCTCATTTGAGAA 2669

Qy 730 LeuLysPheTyrGluAspGlnGlyLeuProAlaLeuLysTyrThrGluLeuAlaGly 749

Db 2670 GAGAGGTTTATGAAGCGCTGCGAGCGCTGTGACTCTCCATCTTCAAGAGGCTAGCTGAC 2729

Qy 750 LeuLeuThrGlnCysMetAlaTyrAspProGlyArgArgProSerPheArgAlaLeuLeu 769

Db 2730 CTCTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2789

Qy 770 ArgAspLeuAsnGlyLeuLeuThrSerAspTyrGluLeuLeuSerAspProThrProGly 789

Db 2790 AGGCACATTAAACCTGAGGAGCAGATCCAGACATTGTTTCAGAA 2837

Qy 790 IleProSerProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAsp 809

Db 2838 ---AAGCAGCCCAACAACAGAGGTG-----GAC 2861

Qy 810 ProAlaIlePheGluGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPhe 829

Db 2862 CCCACTCTCTTGAAGCGCTTCTGAGAGGATTCGTGACTTCGGAGAGGTCACATT 2921

Qy 830 GlySerValGluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAla 849

Db 2922 GGAAGTTGAGCTCTGAGATATGATCTGAGGGAGACACACAGGGGAGCAGGTAGCT 2981

Qy 850 ValLysGlnLeuGln---HisSerValProAspGlnGlnArgAspPheGlnArgGluIle 868

Db 2982 GTCAAGTCCCTGAAGCTGAGAGTGGAGTAAACCATAGCTAGCTGTGAAGAGAGATA 3041

Qy 869 GlnIleLeuLysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGly 888

Db 3042 GAGATCTTACGGAACCTTACCATTGAGAGACATTGTGAAGTACAAAGAACTGTCATGAA 3101

Qy 889 ProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAsp 908

Db 3102 GACGGAGGAAGTGTATCAAGCTCATCATGGAGTTCTGCTTCGGGAAGCCTTAAGAGG 3161

Qy 909 LeuLeuGlnArgHisArgGly---LeuHisThrAspArgLeuLeuLeuPheAlaTyrGln 927

Db 3162 TATCTGCCAAGAATAAGAACAAATCAACCTCAACAGCAGCAGTAAATATGCTCATCCAG 3221

Qy 928 IleCysLysGlyMetGluTyrLeuGlyAlaArgCysValHisArgAspLeuAlaAla 947

Db 3222 ATTGTGAAGGGATGACTACTTGGGTTCTCGCAATACGTTTCCCGGAGCTTAGCAGCA 3281

Qy 948 ArgAsnIleLeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLys 967

Db 3282 AGAATGCTCTTGTGAGAGTGAGCATCAAGTGAAGATCGAGACTTTGGTTTAAACCAA 3341

Qy 968 LeuLeuProLeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerProIlePhe 987

Db 3342 GCAATTGAACCCGATAGGAGTACTACAGTCAAGGACGACCGGGACAGCCAGTGTTTC 3401

Qy 988 TrpTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTyrSer 1007

Db 3402 TGTACCTCTCGGAATGTTTAACTCCAGTGAATTTTATATGCTCTGATGCTGTCT 3461

Qy 1008 PheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAla 1027

Db 3462 TTTGGAGTGACCTGCAGAGCTGCTCATTACTGTGACTCAGATTTTGTCTCCATGGCC 3521

Qy 1028 GluPheLeuArgMetMetGlyProGluArgGluGlyProProLeuCysArgLeuLeuGlu 1047

Db 3522 TTGTCTCTGAAATGATAGGCCCAACTCATGGCCAGATGACGACACCGCTGTGAAG 3581

Qy 1048 LeuLeuAlaGluGlyArgArgLeuProProProProThrCysProThrGluValGlnGlu 1067

Db 3582 ACTCTGAAGAGGAGAGCGTCTGCCATGTCCACCACTGCTCTGATGAGTTTATCAG 3641

Qy 1068 LeuMetGlnLeuCysTrpAlaProGluProHisAspArgProAlaPheAlaThrLeu 1086

Db 3642 CTTATGAGAAATGCTGGGAATTCACCACTCAACCGGACAACTTTTTCAGAACCTT 3698

RESULT 9

US-09-880-107-2379

Sequence 2379, Application US/09880107

Patent No. US20020142981A1

GENERAL INFORMATION:

APPLICANT: Horne, Darci T.

APPLICANT: Vockley, Joseph G.

APPLICANT: Scherf, Uwe

APPLICANT: Gene Logic, Inc.

TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

FILE REFERENCE: 44921-5028-MO

CURRENT APPLICATION NUMBER: US/09/880,107

CURRENT FILING DATE: 2001-06-14

PRIOR APPLICATION NUMBER: US 60/211,379

PRIOR FILING DATE: 2000-06-14

PRIOR APPLICATION NUMBER: US 60/237,054

PRIOR FILING DATE: 2000-10-02

NUMBER OF SEQ ID NOS: 3950

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2379

LENGTH: 3541

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020142981A1 M64174

US-09-880-107-2379

Alignment Scores:

Pred. No.: Length: 3541

Score: 443

Percent Similarity: 53.66%

Best Local Similarity: 37.67%

Query Match: 32,49%

Indels: 164

Gaps: 24

DB:

US-09-397-967A-16 (1-1099) x US-09-880-107-2379 (1-3541)

Qy 17 SerLeuSerSerSerGluAlaGlyAlaLeuHisValLeu---LeuProProArgGlyPro 35

Db 121 AACCTGGAGGCCCTGAGCCAGGGGTGGAAGTATCTTCTATCTGTCGACAGAGGAGCCC 180

Qy 36 GlyProProGlnArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysVal 55

Db 181 -----CTCGGGTGGCAGTGGAGAGTACACAGCAGGAACTGTGCATC 225

Qy 56 ArgAlaAlaLysAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThr 75

Db 226 AGGGCTGCACAGCAGTCCGCTATCTCTCTCTTTGTCAACACTCTTTGCCCTGTATGAC 285

Qy 76 GluAspPheSerCysTyrPheProProSerHisIlePheCysIleGluAspValAspThr 95

Db 286 GAGAACACCAAGCTCTGGTATGCTCCAAATCGCACCATCACCGTTGATGACAGATGTC 345

Qy 96 GlnValLeuValTyrArgLeuArgPheTyrPheProAspTrpPheGlyLeu----- 112

Db 346 CTCGGCTCCACTACCGGATGAGGTTCTATTTCCCAATTGGCATGGACCAACAGCAAT 405

Qy 113 ---GluThrCysHisArgPheGlyLeuArgLysAspLeu----- 124

Db 406 GAGCAGTCACTGTGGCTGCTATCTCCAAAGACAGCAAGAAATGGCTACGAGAAAAAAG 465

Qy 125 -----ThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141

Db 466 ATCCAGATCAACCCCTCTCTTGTATGTCAGCTCACTGGAGTATCTTTGTCTCAGGGA 525

Qy 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGln----- 159

Db 526 CAGTATGATTGGTGAATGCTGCTTATTCGAGCCCAAGACCGCAGCAGGATGGA 585

Qy 160 -----GlyGluPheLeuSerLeuAlaValLeuAspLeuGlnMetAlaArg 175

Db 586 CATGATATTGAGAACAGAGTGTCTAGGATGGCTCTCTGGCCATCTCACACTATGCCATG 645

QY	176	GluGlnAlaGlnArgProGlyGluLeuLeuLeuLysThrValSerTyrLysAlaCysLeuPro	195
DB	646	ATGAGAAGATGCAGTTGCCAGACTGCCAAGGACATCAGCTACAAAGCATATATATCCA	705
QY	196	ProSerLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgArgIle	213
DB	706	GAACATTGAATAAGTCCATCAGACAGAGGAACCTTCTCACCGAGATCGGATAAATAAT	765
QY	214	-----ArgArgThrValValLeuAlaLeuLeuPro	223
DB	766	GTTTTCAAGGATTTCTAAAGGAATTTAAACAACAGACCAATTTGTGACAGAGCGTG	822
QY	224	CysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGlu	243
DB	823	-----TCCACGATCAGCTGAAGGTGAATACTCTGGCTACCTTGGAA	864
QY	244	ArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGlu	263
DB	865	ACTTTGACAAACATTACGGTGCTGAATAATT	897
QY	264	GluProGlyLeuLeuArgValAlaGlyAspIleProTrp---SerSerAsnAsp	282
DB	898	GAGACTTCCATGTTACTGAATTCATCAGAAATGAGATGAATGGTTTCATTCGAATGAC	957
QY	282	-----	282
DB	958	GGTGGAAACGTTCTACTACGAAGTGATGTGTGCTGGGAATCTTGGAAATCCAGTGGAGG	1017
QY	282	-----	282
DB	1018	CATAAACCAAAATGTTGTTCTGTGAAAAAGGAAAAATAAATGAAGCGGAAAAAACTG	1077
QY	283	-----GluLeuPheGlnThrPhe	288
DB	1078	GAAATAAAGACAAGAAGGATGAGGAGAAAAACAAGATCCGGGAAGATGGAAACAATTTT	1137
QY	289	CysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyProAla	308
DB	1138	TCATTCTTCCTGAAATCACTCACTTGTATTAAGGAGTCT	1179
QY	309	GlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPhe	328
DB	1180	-----GTGGTCAGCATTAACAAGCAGGACACAAGAAAAATGGAATCTGAAGCTC	1227
QY	329	ProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIle	348
DB	1228	TCITCCACAGAGAGGCTTGCTCTTGTGTCCCTGTAGATGGCTACTTCCGGCTCACA	1287
QY	349	CysAspSerArgHisTyrPheCysLysGluValAlaProArgLeuLeuGluGlu	368
DB	1288	GCAGATGCCCATCATTACCTCTGCACCGAGTGGCCCCCGTGTGATCGTCCACAACATA	1347
QY	369	AlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAlaAla	388
DB	1348	CAGAATGGCTGTCATGTGTCCAATCTGTACAGAATACGCCATCAATAAATTCGGCAAGAA	1407
QY	389	GlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeu	408
DB	1408	GGAAGCAGAGGGGATGTACGTCTGAGGTGGAGCTGCACCGCACTTTGACACATCCTC	1467
QY	409	LeuThr---AlaCysValGlnThrPro-----LeuGlyProAspTyrLys	422
DB	1468	ATGACCGTCACTGCTTTTGAGAGTCTGACGAGTGCAGGGTGCCTCCAGAGACGAGTTCAG	1527
QY	423	GlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnPro	442
DB	1528	AACATTTTCAGATC---GAGGTGCAAGAGGCGCTACAGTCTGCACGCTTCGGACCGCAGC	1584
QY	443	HisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGly	462
DB	1585	TTCCCCAGCTTGGGACCTCATGAGCCACTCAAGAAGCAGATCTCTCGACGGATAAC	1644

QY	463	AlaAlaLeuTyrLeuThrSerCysCysAlaProArgProIysGluIysSerAsnLeuLeu	482
DB	1645	ATCAGCTTCATGCTAAACCGCTGTCGACGCCAACCCGAGAAATCTCCAACTGCTG	1704
QY	483	ValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysCysAla	502
DB	1705	GTGGCTACTAAG-----AAGCCCGAGGAGTGGCAGCCCC---GTCTACCCC	1746
QY	503	LeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTyrPheHisGluAsnLeu	522
DB	1747	ATGAGCCAGCTGAGTTTCGATCGGATCCTCAAGAAGGATCTGGTGCAGGCGGAGCACCTT	1806
QY	523	GlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValValAspGlyGlu	542
DB	1807	GGGAGAGCGCAGAGAACACACATCTATTCTGGGACCCCTGATGGATTACAAGGATACGAA	1866
QY	543	ThrHisAspSerGlu-----ValLeuLeuLysValMetAspSerArgHisArg	558
DB	1867	GGAACTCTCGAAGAGAAGAGATAAAGTGATCTCTCAAGTCTTTAGACCCCGACCCACAGG	1926
QY	559	AsnCysMetGlnSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHis	578
DB	1927	GATATTTCCCTGGCCCTTCTTCGAGCCAGCAGCATGATGAGACAGGTCTCTCCACAAACAC	1986
QY	579	LeuValLeuLeuHisGlyValCysMetAlaGly---AspSerIleMetValGlnGluPhe	597
DB	1987	ATCGTGATCTTATGCGCGTCTGTGTCGCGCAGCGTGGAGAATATCATGTTGGTGAAGAGTTT	2046
QY	598	ValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSer	617
DB	2047	GTGGAAGGGGTCTCTGGATCTCTTCATGCACCGGAAAGATGATGCTTACCCACCA	2106
QY	618	TrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGly	637
DB	2107	TGGAAATTCAAAGTTGCCAACACAGCTGGCCAGTGCCTGAGCTACTTGGAGGATTAAGAC	2166
QY	638	LeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyIysAspGly	657
DB	2167	CTGTCTCATGGAATGTGTGTACTTAAACCTCTCTCTGGCCGTGAGGGAATCGACAGT	2226
QY	658	Asn---ProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeu	676
DB	2227	GAGTGTGGCCATTTCATCAAGCTCAGTGACCCCGGCATCCCCCAATACGGTGTCTAG	2286
QY	677	GlyMetLeuThrAspArgIleProTyrValAlaProGluCysLeuGlnGluAlaThr	696
DB	2287	CAAGAAATGCATTGAAACGAATCCCATGGATGTCTCTGAGTGTGTGAGAGACTCCAGAAC	2346
QY	697	LeuCysLeuGluAlaAspLysTyrGlyPheGlyAlaThrThrTrpGluValPheGlnArg	716
DB	2347	CTGAGTGTGCTGCTGACAGTGGAGCTTTGGAAACCCACGCTCTGGGAAATCTGCTACAAT	2406
QY	717	GlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGln	736
DB	2407	GGCAGATCCCTTGAAGACAGACGCTGATTGAGAAAGAGAGATTCATGAAAGCCGG	2466
QY	737	GlyGlnLeuProAlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMetAla	756
DB	2467	TGCAGCCGAGTACACCATCATGTAGGAGCTGGCTGACCTCATGACCCGCTGCATGAAC	2526
QY	757	TyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuLeu	776
DB	2527	TATGACCCCAATCAGAGGCGCTTCTTCCGAGCCCATCATGAGAGACATTAATAAGCTTGA	2586
QY	777	ThrSerAspTyrGluLeuLeuSer-----AspProThrProGlyIleProSer	792
DB	2587	GAGCAGAAATCCAGATATTGTTTCCGAAAAAAAACCCAGCCCAACTGAA-----	2634
QY	793	ProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIle	812
DB	2635	-----GTGGACCCCAACAT	2649
QY	813	PheGluGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerVal	832

```

Db 2650 TTTGAGAGCGCTCTTAAGAGAGTCCGTGACCTGGAGAGGCCACTTTGGGAGAGTT 2709
Qy 833 GluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValIysGln 852
Db 2710 GAGCTCTGCAGGTATGACCC--GAAGACAATACAGGGAGCAGGTGGCTGTAAATCT 2766
Qy 853 LeuGln---HisSerValProAspGlnGlnArgAspPheGlnArgGluLeuGlnLeu 871
Db 2767 CTGAAGCTGAGAGTGGAGGTAAACCATAGCTGATCTGAAGAAAGGAATCGAGATCTTA 2826
Qy 872 LysAlaLeuHisSerAspPheIleValIysTyrArgGlyValSerTyrGlyProGlyArg 891
Db 2827 AGGAACCTCTATCATGAGAACATTGTGAAGTACAAAGGAATCTGCACAGAGAGCGGAGA 2886
Qy 892 GlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGln 911
Db 2887 AATGGTATTAAAGCTCATCATGGAATTTCTGCTTCGGGAAGCTTTAAGGAATATCTTCCA 2946
Qy 912 ArgHisArgGly---LeuHisThrAspArgLeuLeuPheAlaTyrGlnIleCysLys 930
Db 2947 AAGAATAAGACAAATAAATACCTCAACAGCAGCTAAATATGCGCTTCAGATTGTGAAG 3006
Qy 931 GlyMetGluTyrLeuGlyAlaArgCysValHisArgAspLeuAlaAlaArgAsnIle 950
Db 3007 GGGATGGACTATTGGGTTCTCGCAATACGTTACCGGACTTGGCAGCAAGAAATGTC 3066
Qy 951 LeuValGluSerGluAlaHisValIysIleAlaAspPheGlyLeuAlaIysLeuPro 970
Db 3067 CTTGTTGAGAGTGAACACCAAGTGAATTTGAGACTTCGGTTTAAACCAAGCAATGAA 3126
Qy 971 LeuGlyLysAspTyrValValArgGluProGlyGlnSerProIlePheTyrAla 990
Db 3127 ACCGATAAGGATATTACACCTCAAGGATGACCGGACAGCCCTGTGTTGGTATGCT 3186
Qy 991 ProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTyrSerPheGlyVal 1010
Db 3187 CCAGATGTTTAAATCAATCTAAATTTATATTATGCTCTGACGCTGGTCTTTGGAGTC 3246
Qy 1011 ValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeu 1030
Db 3247 ACTCTGCATGAGCTGCTGACTTACTGTGATTGAGATTCAGATTCCTAGTCCATGGCTTTGTCCTG 3306
Qy 1031 ArgMetMetGlyProGluArgGluGlyProProLeuCysArgLeuLeuLeuAla 1050
Db 3307 AAAATGATAGGCCCAACCCATGGCCAGATGACAGTCACAGACTGTGTAATACGTAAAA 3366
Qy 1051 GluGlyArgArgLeuProProProProThrCysProThrGluValGlnGluLeuMetGln 1070
Db 3367 GAAGGAAACGCTGCGTGGCCACCTAACTGTCCAGATGAGGTTTATCAGCTTATGAGA 3426
Qy 1071 LeuCysTyrAlaProGluProHisAspArgProAlaPheAlaThrLeu 1086
Db 3427 AAATGCTGGAAATTCCAACCATCCATCGGACAAGCTTTTCAAGAACCTT 3474

```

RESULT 10

```

US-09-873-367C-163
; Sequence 163, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-84
; CURRENT APPLICATION NUMBER: US/09/873,367C
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29

```

```

; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 163
; LENGTH: 3541
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-873-367C-163

Alignment Scores:
Pred. No.: 6,36e-186 Length: 3541
Score: 1904.00 Matches: 443
Percent Similarity: 53.66% Conservative: 188
Best Local Similarity: 37.67% Mismatches: 381
Query Match: 32.49% Indels: 164
DB: 13 Gaps: 24

US-09-397-967A-16 (1-1099) x US-09-873-367C-163 (1-3541)
Qy 17 SerLeuSerSerSerGluAlaGlyAlaLeuHisValLeu---LeuProProArgGlyPro 35
Db 121 AACCTGGAGGCCCTCAGCCAGGGGTGGAAGTATCTTCTATCTGCGGACAGGAGGCC 180
Qy 36 GlyProProGlnArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysVal 55
Db 181 -----CTCCGGCTGGCGAGTGGAGAGTACACAGCAGAGGAACCTGTGCATC 225
Qy 56 ArgAlaAlaLysAlaCysGlyLeuProValTyrHisSerLeuPheAlaLeuAlaThr 75
Db 226 AGGGCTGACAGGCATGCGTATCTCTCTCTTTGCACAACTCTTTGCCCTGTATGAC 285
Qy 76 GluAspPheSerCysTyrPheProProSerHisIlePheCysIleGluAspValAspThr 95
Db 286 GAGAACACCAAGCTCTGGTATGCTCCAAATCCACCATCCACCATGATCACAAGATGTC 345
Qy 96 GlnValLeuValTyrArgLeuArgPheTyrPheProAspTyrPheGlyLeu----- 112
Db 346 CTCCTGGCTCCACTACCGGATGAGGTTCTATTTCACCAATTGGCATGGAACCAACGACAT 405
Qy 113 ---GluThrCysHisArgPheGlyLeuArgLysAspLeu----- 124
Db 406 GAGCAGTCAGTGGCGTCATTCTCCAAAGACGACAAAATGGCTACGAGAAAAAAG 465
Qy 125 -----ThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141
Db 466 ATTCAGATGACCAACCCCTCTCTGATGCCAGTCACTGAGATATCTGTGTCAGGGA 525
Qy 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGln----- 159
Db 526 CAGTATGATTGGTGAATGCTGCTGCTCTATTTCGAGACCCCAAGACCCGAGGATGGA 585
Qy 160 -----GlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArg 175
Db 586 CATGATATTGAGAACAGTGTCTAGGATGGCTGTCTGCGCATCTCACACTATGCCATG 645
Qy 176 GluGlnAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrIysAlaCysLeuPro 195
Db 646 ATGAAGAAGATGCAGTTGCCAGAACTGCCCAAGGACATCAGCTTACAGGATATATTCCA 705
Qy 196 ProSerLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgIle----- 213
Db 706 GAACATTTGAATGAATGCTCATCAGACAGAGGAGCACTTCTCACCAGGATGGGATTAAT 765
Qy 214 -----ArgArgThrValValLeuAlaLeuLeuPro 223
Db 766 GTTTTCAAGGATTTCTCTAAAGGAATTTAAACACACAGACCATTTGTGACAGCAGGCTG 822
Qy 224 CysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGlu 243

```

Db 823 -----TTCACGCGATACCTGGAAGGTGAATACTTCGCTACCTTGGAA 864
Qy 244 ArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGlu 263
Db 865 ACTTGGACAAACATTACGGTGCTGAATATTT----- 897
Qy 264 GluProGlyLeuLeuArgValAlaGlyAspAenGlyIleProTirP---SerSerAsnAsp 282
Db 898 GAGACTTCATGTTACTGATTTTCACAGAAATGAGATGAATGGTTTCATTTCGNATGAC 957
Qy 282 ----- 282
Db 958 GGTGGAAACGTTCTCTACTACGAAGTGTGCTGACTGGGAATCTTGGAACTCCAGTGGAGG 1017
Qy 282 ----- 282
Db 1018 CATAAACCAATGTGTTCTGTTGAAAGGAAATAAATAACAGCGGAAAAAATCG 1077
Qy 283 -----GluLeuPheGlnThrPhe 288
Db 1078 GAAATAAAGACAAAGAGTACGAGAGAAAACAAGATCCCGGAAGAGTGGAAACATTTT 1137
Qy 289 CysaspPheProGluIleValaspValSerIleAsnGlnAlaProArgValGlyProAla 308
Db 1138 TCATCTTCCTCGAATCACTCACTTGTATTAAGAGGAGTCT----- 1179
Qy 309 GlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuLeuGluAlaGluPhe 328
Db 1180 -----GTGGTTCAGATTAAACAGCAGGACCAACAGAAATATGGAACTCAAGCTC 1227
Qy 329 ProGlyLeuProGluAlaLeuSerPheValAlaLeuValaspGlyTyrPheArgLeuIle 348
Db 1228 TCTTCCACAGAGGCGCTTGCTTGTGTCCTGCTAGATGCTACTTCCGGCTCACA 1287
Qy 349 CysaspSerArgHisTyrPheCysLeuValAlaProProArgLeuGluGluGlu 368
Db 1288 GCAGATGCCCATATTACCTCTGCACCGCGTGGCCCGCCCGTGTGATCGTCCACACATA 1347
Qy 369 AlaaspValCysHisGlyProIleThrLeuaspPheAlaIleHisTysLeuLysAlaAla 388
Db 1348 CAGATGGCTGTATGTCTCAATCTGTACAGATACCGCATCAATAAATGGCGCAAGAA 1407
Qy 389 GlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeu 408
Db 1408 GGAAGCGAGAGGGGATGTACGTGCTGAGTGGAGCTGCACCGACCTTTGACACATCCTC 1467
Qy 409 LeuThr---AlaCysValGlnThrPro-----LeuGlyProAspTyrLys 422
Db 1468 ATGACCGTCACTGCTTTGAGAAGTCTGACAGGTGCAGGCTGCCAGAGCAGTTCAG 1527
Qy 423 GlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnPro 442
Db 1528 AACTTTCAGATC---GAGGTGGAGAGGGCGGCTACAGTCTGCAGGCTTCGGACCGCAGC 1584
Qy 443 HisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValaspGly 462
Db 1585 TTCCCGCAGCTTGGACACCTCATGAGCCACCTCAAGAGCAGATCTTCGCCACGGATAC 1644
Qy 463 AlaAlaLeuTyrIleuThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIle 482
Db 1645 ATCAGCTTCATGTAAACCGCTGCTGCCAGCCCAAGCCCGAGAAATCTCCAACCTGCTG 1704
Qy 483 ValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysCysAla 502
Db 1705 GTGGCTACTAAG-----AAAGCCAGGAGTGGCAGGCC---GTCTACCCC 1746
Qy 503 LeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluThrPheGluAsnLeu 522
Db 1747 ATGAGCCAGCTGAGTTTCGATCGATCCTCAAGAGGATCTGGTGCAGGCGGAGCACCTT 1806
Qy 523 GlyHisGlySerPheThrLysIlePheArggGlySerArgGluValValaspGlyGlu 542
Db -----

Db 1807 GGGAGAGGACGAGAACACACATCTATTCTGGGACCTGATGATTAACAGGATGACGAA 1866
Qy 543 ThrHisAspSerGlu-----ValLeuLeuLysValMetAspSerArgHisArg 558
Db 1867 GGAACCTTCTGAAGAGAGAGATAAAGTGTCTCAAAAGTCTTAGACCCAGCCACAGG 1926
Qy 559 AsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHis 578
Db 1927 GATATTTCCCTCGGCTTCTTCGAGGACCCAGCAGCATGTAGACACAGGTCTCCACAAAC 1986
Qy 579 LeuValLeuLeuHisGlyValCysMetAlaGly---AspSerIleMetValGlnGluPhe 597
Db 1987 ATCGTGTACCTCTATGGGCTGTGTGCGCAGCTGGAGAAATATCATGTGGAAGAGTTT 2046
Qy 598 ValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSer 617
Db 2047 GTGGAAGGGGCTCTCTGGATCTCTCATGCACCGGAAAGTGTATGCTCTTACCACACCA 2106
Qy 618 TrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGly 637
Db 2107 TGGAAATTCAAAGTTGCCAACACAGCTGGCCAGTCCCTGAGCTACTTTGGAGATAAAGAC 2166
Qy 638 LeuProHisGlyAsnValSerAlaArgLysValLeuAlaArgGluGlyGlyAspGly 657
Db 2167 CTGCTCCATGGAATGTGTACTAAACCTCTCTGCGCCGTGAGGGAATCGACAGT 2226
Qy 658 Asn---ProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeu 676
Db 2227 GAGTGTGGCCCATTCATCAAGCTCAGTGACCCCGCATCCCATTCATCGTGTCTCTAGG 2286
Qy 677 GluMetLeuThrAspArgIleProTyrValAlaProGluCysLeuGlnGluAlaGlnThr 696
Db 2287 CAGAATGCAATGCAACGAATCCCATGATGCTCTGAGTGTGTGAGGATCCCAAGAAC 2346
Qy 697 LeuCysLeuGluAlaAspLysTyrGlyPheGlyAlaThrThrTrpGluValPheGlnArg 716
Db 2347 CTGAGTGTGCTGCTGCACAGTGGAGCTTTGGAAACACCGCTCTGCGAAATCTGCTACAAT 2406
Qy 717 GlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGln 736
Db 2407 GGGCAGATCCCTTGAAGACAAAGACGCTGATTGAGAAAGAGAGATTCATGAAAGCCGG 2466
Qy 737 GlyGlnLeuProAlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMetAla 756
Db 2467 TGCAGCCAGTGCACACCATCATTAAGAGCTGGCTGACCTCATGACCCGCTGCATGAAC 2526
Qy 757 TyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIle 776
Db 2527 TATGACCCCAATCAGAGGCTTTCTTCGAGGCCATCATGAGACATTAATAAGCTTGAA 2586
Qy 777 ThrSerAspTyrGluLeuLeuSer-----AspProThrProGlyIleProSer 792
Db 2587 GAGCAGAATCCAGATATGTTTCCAGAAAAAACCAGCAACTGAA----- 2634
Qy 793 ProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIle 812
Db 2635 -----GTGGACCCCCACACAT 2649
Qy 813 PheGluGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerVal 832
Db 2650 TTTGAGACCGCTTCTTAAGAGGATCCGTGACTTTGGAGAGGGCCACTTTGGGAAGGTT 2709
Qy 833 GluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGln 852
Db 2710 GAGCTCTGCAGGTATGACCCC---GAAGACATACAGGGGAGCAGGTGCTGTAAATCT 2766
Qy 853 LeuGln---HisSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnLeu 871
Db 2767 CTGAAGCCCTGAGAGTGGAGGTAAACACATAGTGTATCTGAAAAAGAAATCGAGATCTTA 2826
Qy 872 LysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArg 891
Db 2827 AGGAACCTCTATCATGAGAACATTGTGAAGTAAAGGAATCTGCACAGAACGCGAGGA 2886

Db 3196 ATGCAATCAATTTTATTTATTTGCTCTGACGCTGCTGCTTTTGGAGTCACTCTGCAATGAG 3255
Qy 1015 LeuPheThrTyrcysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034
Db 3256 CTGCTGACTTACTGTGATTTCAGATTCTAGTCCCATGGCTTTGTTCTCTGAAATGATAGGC 3315
Qy 1035 ProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArg 1054
Db 3316 CCAACCCATGCCAGATGACAGTCAAGACTTGTGAATACGTTTAAAGAGAGAAACGC 3375
Qy 1055 LeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAla 1074
Db 3376 CTGCGTGCCACCTTAACGTCTCCAGATGAGGTTTATCACTTATGAGGAATGCTGGGAA 3435
Qy 1075 ProGluProHisAspArgProAlaPheAlaThrLeu 1086
Db 3436 TTCCAACCATCAATCGGACAGCTTTCAGAACCTT 3471

RESULT 12
US-10-305-720-1353
; Sequence 1353, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1353
; LENGTH: 4080
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: GenBank ID No. US20040010136A1 g37503
US-10-305-720-1353

Alignment Scores:
Pred. No.: 6, 8e-177 Length: 4080
Score: 1817.50 Matches: 464
Percent Similarity: 50.49% Conservative: 151
Best Local Similarity: 38.10% Mismatches: 432
Query Match: 31.02% Indels: 171
DB: 12 Gaps: 30

US-09-397-967a-16 (1-1099) x US-10-305-720-1353 (1-4080)

Qy 4 ProSerGluGluThrProLeu-----IleProGlnArgSerCysSerLeuSer----- 19
Db 280 CCCTCAGCTGAGCCACCTCGCGGAGCATGCTCTCGCCACTCGGGGATGGCGGCG 339
Qy 20 -----SerSerGluAlaGlyAlaLeuHisValLeu 29
Db 340 AGTAAGCCCTTGGGATGAGCCGACCATGCTGCCATGGGAGCCCTGAAGTGCTT 399
Qy 30 LeuProProArgGlyProGlyProProGlnArg---LeuSerPheSerPheGlyAspTyr 48
Db 400 CTGCACTGGGCTGCTCCAGCGGGGAGCCCTCGGTCACTTTCAGTGAGTCATCGCTG 459
Qy 49 LeuAlaGluAspLeuCysValArgAlaAlaLysAlaCysGlyIleLeuProValTyrHis 68
Db 460 ACAGCTGAGGAAGTCTGCATCCCATTTGCATAAAGTTGGTATCACTCCTCTGCTTC 519
Qy 69 SerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPheProSerHisIlePhe 88
Db 520 AATCTCTTTGCCCTCTTTCGATGCTCAGGCCCAAGTCTGGTGTGCCCCCAACCATCTTA 579
Qy 89 CysIleGluAspValAspThrGlnValLeuValTyrArgLeuArgPheTyrPheProAsp 108
Db 1111 CGCTTCGCGCACAGAGCGTGTGCCCTGTGCCCTGTGCCCGGCGGAGGG 1170

Db 580 GAGATCCCGACAGAGATGCAAGCCTGATGCTATATTTCCGCATAGAAGTTTTTTTCCGGAAC 639
Qy 109 TrpPheGlyLeuGlu-----ThrCysHisArgPheGly----- 119
Db 640 TGGCATGCGATGAATCCTCGGAAACCGGCTGTGTACCGTTTGTGGGCCCCCAGGAACCGAG 699
Qy 120 LeuArgLysAspLeuThrSer-----AlaIleLeuAspLeuHisValLeuGluHis 136
Db 700 GCATCCTCAGATCAGACAGCAGGGGATGCAACTCTCTGGACCCAGCCTCATTTGAGTAC 759
Qy 137 LeuPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProVal---GlyLeuSer 155
Db 760 CTCTTTGAGCAGGCAAGCATGAGTTTGTGAATGACGTGGGCATCACTGTGGGAGCTGTGCG 819
Qy 156 MetLysGluGln-----GlyGluPheLeuSerLeuAlaValLeuAspLeu 170
Db 820 ACCGAGGAGGAGATCCACACTTTAAGATGAGAGCCTGGGCATGGCCTTTCTGCACCTC 879
Qy 171 AlaGlnMetAlaArgGluGlnAlaGlnArgProGlyGluLeuLeuLysThrValSerTyr 190
Db 880 TGTCACTCGCTCTCCGCCATCGCATCCCTGTGAGAGGAGTGGCCCAAGAACAGCAGCTTC 939
Qy 191 LysAlaCysLeuProProSerLeuArgAspValIleGlnGlyGlnAsnPheValThrArg 210
Db 940 AAGGACTGCATCCCGCGCTCTTCCCGCGCATATCCGCGCACACAGCCCTGACCCCG 999
Qy 211 ArgArgIleArg-----ArgThrValValLeuAlaLeuLeuProCysGlyArgLeu 227
Db 1000 CTGCGCTTCGGAAGCTCTTCGCGAGTTCCTCGGGGACTTCCAGCGC---GGCCGACTC 1056
Qy 228 ProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGluArgLeuHisPro 247
Db 1057 TCCAGCAG-----ATGCTCATGTCAAATACCTAGCCACACTCGAGCGGTGGCACCC 1110
Qy 248 AlaAlaThrThrGluThrPheArgValGly-----LeuProGlyAlaGlnGlu 263
Db 1111 CGCTTCGCGCACAGAGCGTGTGCCCTGTGCCCTGTGCCCGGCGGAGGG 1170
Qy 264 GluProGlyLeuLeuArg----- 269
Db 1171 GAGCCCTGTATCATCCGGACAGTGGGTGGCGCCCTACAGACCTGGCCCTGAGTCTGCT 1230
Qy 270 -----ValAlaGlyAspAsnGlyIleProThrSerSer 280
Db 1231 GCTGGGCCCCCAACCCACGAGGTGCTGTGTGACAGGCACTGTGTGTCATCCAGTGTGGCCA 1290
Qy 281 AsnAspGlu----- 283
Db 1291 GTAGAGGAGGAGTGAACAAGGAGGAGGTTCGTGTGGCAGCAGTGGCAGGAACCCCCAA 1350
Qy 283 ----- 283
Db 1351 GCCAGCCTGTTGGGAAGAGCCAGAGCTCACAGGCATTCGCCCGAGCGGCGAGCAGG 1410
Qy 284 -----LeuPheGlnThrPheCysAspPheProGluIleValAspValSerIle 299
Db 1411 CCGCGGAGGCCACTGTGGGCGCTACTTCTGTGACTTCGGGACATCACCCACGCTGTGCTG 1470
Qy 300 AsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMet 319
Db 1471 AAA-----GAGCAC---TGTGTGAGCATCCACCGCGAG 1500
Qy 320 AspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAla 339
Db 1501 GACCAAGTGCCTGGAGCTGAGCTTGCCTTCCCGGGCTCGCGCGCTGTCTTCTGTGTCG 1560
Qy 340 LeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluVal 359
Db 1561 CTGTTGAGCGGCTATTTCGCGCTGAGCGCGCATCTCCAGCCACTACCTGTGCCACGAGGTG 1620
Qy 360 AlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIleThrLeuAsp 379
Db 1621 GCTCCCCACGCGTGTGTGATGAGCATCCGGGTGGGATCCACGAGCCCTGCTGGAGCCA 1680

QY	714	PheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaLysLeuLysPheTyr	733
DB	2728	TGCTTTGACGGAGAGGCCCTCTTCGAGAGCGCAGTCCCTCCGAGAAAGGAGCATTTTCTAC	2787
QY	734	GluAspGlnGlyGlnLeuProAlaLeuLysTrpThrGlnLeuAlaGlyLeuIleThrGln	753
DB	2788	CAGAGGCGACCGGCTGCCGAGCCCTCTGCCCCACAGCTGGCGCACACTCACCGACGAC	2847
QY	754	CysMetalTyraAspProGlyArgProSerPheArgAlaIleLeuArgAspLeuAsn	773
DB	2848	TGCTGACTATGACCAACCCAGAGGCCATCATTCGCGACCATCTCTGCGTGGACCTCACC	2907
QY	774	GlyLeuIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerPro	793
DB	2908	CGCGTG-----CAGCCCCCAATCTTGTGACGCTGTGACTGTGAACGGGACTCACCG	2961
QY	794	ArgAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePhe	813
DB	2962	-----GCGCTCGGACCTACTACTTTC	2982
QY	814	GluGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGlu	833
DB	2983	CACAAGCGCTATTGAAAAAGATCCGAGACTCGGCGAGGGTCACTTCGCGAAGGTACG	3042
QY	834	LeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeu	853
DB	3043	TTGTACTGCTACGATCCGACCAACGACGCGACTCGGCGAGATGCTGGCGTGAAAGCCCTC	3102
QY	854	GlnHisSerValProAspGlnGlnArgAsp---PheGlnArgGluIleGlnIleLeuLys	872
DB	3103	AAGCGACTCGGCCCCCAGCACCGCTCGGGCTGGAAAGCAGGAGATTGACATTCTGGCG	3162
QY	873	AlaLeuHisSerPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGln	892
DB	3163	ACGCTCTACCAGGACACATCATCAGTACAGGGGTGCTGGCAGGACCAAGCGCAGAGAG	3222
QY	893	SerLeuArgIleValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArg	912
DB	3223	TCGCTGCAGCTGTGATGGAGTAGTACGTGCCCTCGGCAGCGCTCCGAGACTACCTGCCCCG	3282
QY	913	HisArgGlyLeuHisThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysGlyMet	932
DB	3283	CAC---AGCATCGGGCTGGCCCCAGCTGCTGCTCTTCGCCACAGCAGATCTCGCAGGGCATG	3339
QY	933	GluTyrLeuGlyAlaArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuVal	952
DB	3340	GCCTATCTGCACGCGCAGCATTACATCCAGCAGACTAGCGCGCGGACACGTGTGTGCTG	3399
QY	953	GluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGly	972
DB	3400	GACAAACAGAGGCTGTGTCAGATCGGGGACTTTGGCTTACGCCAAGCGCGTCCCGAAGGC	3459
QY	973	LysAspTyrTyrValValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGlu	992
DB	3460	CACGAGTACTACCGGTGCGCAGGATGGGGACACCCCGGTGTTCTGGTATGTCGCCACAG	3519
QY	993	SerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeu	1012
DB	3520	TGCCTGAAGGATGTAAGTTCTACTACGTGAGATGTCTGGTCTCTCGGGGTGACCCCTG	3579
QY	1013	TyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMet	1032
DB	3580	TATGAGCTGCTCAGCGACTGTGACTCCAGCCAGAGCCCCCCCCACGAAATTCCTTGAGCTC	3639
QY	1033	MetGlyProGluArgGluGlyProProLeuCysArgLeuLeuLeuLeuAlaGluGly	1052
DB	3640	ATAGGCATTGCTCAGGGTCAGATGACGTTCTGAGACTCACTGAGTTGCTGGAAACGAGGG	3699
QY	1053	ArgArgLeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCys	1072
DB	3700	GAGAGGCTGCCACGGCCCCCAAAATGTCCTGTGAGGTCATCATCTCATGAAGAACTGC	3759
QY	1073	TrpAlaProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeu	1090

Db 1847 AGTCTGCAGGCTCTCGGACCGGAGCTTCCCGACATGGGAGACCTCATGAGCCACCTCAA 1906
Qy 454 pAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProAr 474
Db 1907 GAAGCAGATCTTCGCGCAGGATAACATCAGCTTCATGCTAAACCGTGTGCGAGCCCAA 1966
Qy 474 gProLysGlnLysSerAsnLeuLeuValValArgArgGlyCysAsnProAlaProAlaPr 494
Db 1967 GCCCGGAGAAATCTCCAACTCTGTGTGCTACTAAG-----AAAGCCCA 2011
Qy 494 oGlyCysSerProSerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAs 514
Db 2012 GGATGGGAGAGCC-----GTCTACCCCATGAGCAGCTGAGTTCGATCGCATCTCAAGAA 2068
Qy 514 pSerLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySe 534
Db 2069 GGATCTGTGCGAGGCGAGCACCTTTGGGAGAGCGACGAGAACACATCTATTCTTGGGAC 2128
Qy 534 rArgArgGluValValAspGlyGluThrHisAspSerGlu-----ValLeuLe 550
Db 2129 CCTGATGGATTACAGGATGACGAAGAACTTCTGAAGAGAAAGATAAAAGTGATCTT 2188
Qy 550 uLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAla-AlaSerL 570
Db 2189 CAAGTCTTAGACCCAGCCAGCGGATATTTCCTGCGCTTCTTCGAGGGAAGCCAGCA 2248
Qy 570 eMetSerGlnValSerTyrProHisLeuVal-LeuLeuHisGlyVal-CysMetAlaGl 589
Db 2249 TGATGAGACAGTCTCCACAAAACATCGTGNCCCTCTATGGGCTCTTGTGTCGCGGA 2308
Qy 589 y---AspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuAr 608
Db 2309 CGTGAGAAATATCATGTGTGAAGAGTTGTGGAAGGGGGTCTCTCGGATCTCTTCATGCA 2368
Qy 608 gLysArgGlyHisLeuValSerAlaSerTrpLysLeuGln-ValThrLysGlnLeuAlaT 628
Db 2369 CCGAAAGCCGATGTCTTTACCAACCATGGCAATTCAAAGTGGCGAAACAGCTGGCCA 2428
Qy 628 yAlaLeuAsnTyrLeuGluAspLysGlyLeuProHis-GlyAsnValSerAlaArgLys 647
Db 2429 GTGCCCTGAGCTACTTGAGAGATAAGACCTGGTCCATGGAAATGTGTACTATAAAAC 2488
Qy 648 Val-LeuLeuAlaArgGluGly-GlyAspGlyAsn---ProPro-PheIleLysLeuSer 665
Db 2489 CTTCCTCTGCGCGCTGAGGCGATCCGACAGTGTGGGCCCATTTTCATCAAGCTCAGT 2548
Qy 666 AspProGlyValSerProThrValLeuSerLeuGluMetLeuThrAspArgIleProTrp 685
Db 2549 GACCCCGGATCCCATCTACGCTGTCTAGGCAAGATGCAATTGAACGAATCCCATGG 2608
Qy 686 ValAlaProGluCys-----LeuGlnGluAlaGlnThrLeuCysLeuGluAlaAsp 702
Db 2609 ATTGCTCTGAGTGTGTGAGGACTCCAGAACCTGAGTGGCTGCGTG-----ACA 2662
Qy 703 LysTrpGlyPheGlyAlaThrTrpGluValPheGlnArgGlyProAlaHisIleThr 722
Db 2663 AGTGGAGCTTTGGAAACCACTGCGGAAATCTGTACAAATGCGAGATCCCTTGAAA 2722
Qy 723 SerLeuGluPro-AlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeuProAlaLe 742
Db 2723 AGCAAGACCTGATTGAGAAAGAGATTTCTATGAAGCGGCTGCGAGCGCATGACACC 2782
Qy 742 uLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGlyArgAr 762
Db 2783 ATCATGTAAAGAGCTGTGACCTGACCTGACCCGCTGATGAATATGACCCCAATCAG 2842
Qy 762 gProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyrGluLe 782
Db 2843 GCCTTTCTCCGAGCCATCATGAGACATTAATAGCTTGAAGACAGAAATCCAGATAT 2902
Qy 782 uLeuSer-----AspProThrProGlyIleProSerProArgAspGluLeuCy 798
Db 2903 TGTTCCAGAAAAAACCAGGCAACTGAA----- 2933

Qy 798 sValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLe 818
Db 2934 -----GTGACCCACACATTTTGAAGCGCTCTCT 2965
Qy 818 uLysTyrIleSerLeuLeuGlyLysGlyAsn-PheGlySerValGlnLeuCysArgTyrA 838
Db 2966 AAAGAGGATCCGTGACTTTGGAGAGAGCGCCACCTTTGGGAAGGTTGAGCTCTCGAGTATG 3025
Qy 838 sProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGln----- 854
Db 3026 ACCCCGAAAGGACAAATACAGGGAGCAGGTGGCTGTATAATCTCTGGAGGCTGAGACT 3085
Qy 855 -----HisSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleL 871
Db 3086 GGAGGTAACCATAGC-----TGATCTGCAAAAAGGAAATCGAGGATC 3130
Qy 871 eUlys-----AlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGlyP 889
Db 3131 TTAAGGGAACTCTATTTCATGAGAAC-----ATTGGAAGTACAAAGGAATCTGTACAGAA 3187
Qy 889 rGlyArgGln--SerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAs 908
Db 3188 GACGGAGGAATGGTATTAAAGCTCATCATGAATTTCTGCTTCGGAAGCCTTAAGA 3247
Qy 908 pLeuLeuGlnArgHisArgGly---LeuHisThrAspArgLeuLeuPheAlaTrpGl 927
Db 3248 ATATCTTCCAAAGAAATAAGAACAAATAAACCTCAACAGCAGCTAAAAATATGCGTTCA 3307
Qy 927 nileCysLysGlyMetGluTyrLeuGlyValArgArgCysValHisArgAspLeuAlaAl 947
Db 3308 GATTTGAAGGGATGAGCTATTGGGTCTCGGCAATACCTTACCGGAGCTTGGCAGC 3367
Qy 947 aArgAsnIle-LeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaL 967
Db 3368 AAGAAATGTCTCTGTTGAGAGTGAACACCAAGTGAAATTTGGAGACTTCGTTTAACCA 3427
Qy 967 yLeuLeuProLeuGlyLysAspTyrTyrValVal-ArgGluProGlyGlnSerProile 986
Db 3428 AAGCAATTTGAAACCGCATAGAGATTTATACCCGTACAAGGATGACCGGACAGCCTGTG 3487
Qy 987 PheTrpTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSer-AspValTr 1006
Db 3488 TTTTGGTATGCTCCAGATGTTTAAATGCAATCTAAATTTTATAATGCTCTCTGACGCTG 3547
Qy 1006 p-SerPheGlyValValLeuTyrGluLeu-PheThrTyrCysAspLysSerCysSerPro 1025
Db 3548 GGTCTTTGGAGTCACTCTGCATGAGCTGGCTGACTTACTGTGATTGATTGATTCTAGTCCC 3607
Qy 1026 SerAlaGluPheLeuArgMetMetGlyProGluArgGluGlyProProLeuCysArgLeu 1045
Db 3608 ATGGCTTTGTTCTGTAATAATGATAGGCCCAACCCATGGCCAGATGACAGTCAAGACTT 3667
Qy 1046 LeuGluLeuLeuAlaGluGlyArgArgLeuProProProPro-ThrCysProThrGluVa 1065
Db 3668 GTGAATAGCTTAAAGAGGAAACCGCTCCCGTGGCCACCTAACTGATCCAGATGAGGT 3727
Qy 1065 lGlnGluLeuMetGlnLeuCys-TripAlaProGluProHisAspArgProAlaPheAlaT 1085
Db 3728 TGATCAACTTATGAGGAATGCTTTGGGAATTCACCACTCAATCGGACAGCTTTTCA 3787
Qy 1085 hrLeu 1086
Db 3788 ACCTT 3792

RESULT 14
US-10-205-219-2
; Sequence 2, Application US/10205219
; Publication No. US20030138803A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair

APPLICANT: Brooksbank, Robert
APPLICANT: Pimock, Robert
TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
FILE REFERENCE: WL-A-018200
CURRENT APPLICATION NUMBER: US/10/205,219
CURRENT FILING DATE: 2002-07-24
PRIOR APPLICATION NUMBER: GB 0118354.0
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 197
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2597
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Janus protein tyrosine kinase 1
US-10-205-219-2

Alignment Scores:
Pred. No.: 8,73e-122 Length: 2597
Score: 1281.50 Matches: 315
Percent Similarity: 52.54% Conservative: 150
Best Local Similarity: 35.59% Mismatches: 303
Query Match: 21.87% Indels: 117
DB: 13 Gaps: 19

US-09-397-967A-16 (1-1099) x US-10-205-219-2 (1-2597)
QY 128 IleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHisArgSerAspLeuValSer 147
DB 61 CTCCTGTGATGCCAGTTCAGTGGAGTATCTGTTGGCGAGGACAGTATGATGATCAAA 120
QY 148 GlyArgLeuProValGlyLeuSerMetLysGluGln-----GlyGlu 161
DB 121 TGCCTGGCTCCCATCGGACCCCAAGACAGACAGATGACATGATGAGATGAG 180
QY 162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgPro 181
DB 181 TGCCTGGGATGGCTGTCTGGCCATCTCACATATGCCATGATGAGAGATGCGAGTTG 240
QY 182 GlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProSerLeuArgAspVal 201
DB 241 CCAGAGCTTCCCAAGACATCAGTCAACGCGATATATCCAGAAACATTGTAATAATCC 300
QY 202 IleGlnGlyGlnAsnPheValThrArgArgArgIle----- 213
DB 301 ATCAGACAGAGGAATCTTCTTACAGATGCGAATAAATAATGTTTCAAGAGATTCTTG 360
QY 214 -----ArgArgThrValValLeuAlaLeuLeuProCysGlyArgLeuProGly 229
DB 361 AAGGAATTTAAACACAGACATCTGTGACAGCAGTGTG----- 399
QY 230 ArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGluArgLeuHisProAlaAla 249
DB 400 TCTACACACAGCCTGAAGGTGAGTACCTGCTACCTTGGAAACCTTAACCAAGCAATTAT 459
QY 250 ThrThrGluThrPheArgValGly-----LeuProGlyAlaGlnGluProGlyLeu 267
DB 460 GGAGCCGAGATATTGAGACTTCTATGCTCTCTGATTTTCACAGAAAATGAGCGAGTCGG 519
QY 268 -----LeuArgValAlaGlyAspAsnGly 275
DB 520 TGCCATTTCCACGACGATGGCATGTCTCTATGAAGTCATGTTAAGCAAGAAATCTCGG 579
QY 276 IleProTyrPheSerAsnAsp----- 282
DB 580 ATCCAGTGGCGGACAGAACCAATGTTCTTCTGTTGAAAAGGAAAAATAAATACTGAAG 639
QY 283 -----Glu 283
DB 640 CGGAAAAAACTGGACTATATTAAACACAGAGAGGATGCCAGAAACAACTCCGGGAA 699
QY 284 LeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaPro 303

DB 700 GAGTGGAAACAGTTTTTCTATTTCCCGAGATTACCCACATTTGTTATCAAGAGTCT--- 756
QY 304 ArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIle 323
DB 757 -----GTGGTCAGCATTAACAACAGGACACAAAGAAC 789
QY 324 LeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGly 343
DB 790 ATGGAACATAAAGCTCTCTCTCATGAGGAGCGCTTGTCTTCTGTCCTCGTAGATGCG 849
QY 344 TyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProArg 363
DB 850 TACTTCCGGCTCACTGCGATGCTCACCATTACCTCTGTACTGTGTGCTCCCCGCTG 909
QY 364 LeuLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHis 383
DB 910 ATTGTCCACACATACAGAAATGGCTGTACGGTCCAACTGCACAGATATTGCAATCAAT 969
QY 384 LysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAsp 403
DB 970 AAGCTGGCGCAGGAGGAGTGGAGGGGATGTACGTCTGAGGTGGAGTGCACCGAC 1029
QY 404 TyrAspSerPheLeuLeu-----ThrAlaCysValGlnThrPro-----LeuGly 418
DB 1030 TTTTGACACATTTCTTATTGCTGTCTGCTCGAAGAGTCCGAGGTATTGGGTAAATCAG 1089
QY 419 ProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGly 438
DB 1090 AAGCAGTTCCAGAACTTTCAGATT---GAGGTGCAGAGGGCGCTACAGCTGCATGCG 1146
QY 439 LeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTyrPheAsnSerGlyLeu 458
DB 1147 TCTGTTGACCACTTTCCCGCTGAGAGACCTCATGAACCACTCAAGAAGCAGATCTCG 1206
QY 459 ArgValAspGlyAlaAlaLeuThrSerCysAlaProArgProLysGluLys 478
DB 1207 CGCACGCAATATAGCTTTGTCTGAAACGCTGTCTGCTGAGTCAAGCTCCGAGAAATC 1266
QY 479 SerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerPro 498
DB 1267 TCCAACTCTAGTACCCACTAAG-----AAAGCCAGGAGTGGCAGCCT 1311
QY 499 SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrp 518
DB 1312 ---GTCTACTCTATGAGCCAGCTGAGTTTCGATCGGATCTCTTAAGAAGATATTATCCAA 1368
QY 519 HisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgGluVal 538
DB 1369 GGTGAGCACCTTGGCAGGACACAGAACACATATCTATTCTGGGACCTCTGGGTATAC 1428
QY 539 ValAspGlyGluThrHisAspSerGlu-----ValLeuLeuLysValMetAsp 554
DB 1429 AAGGACATGAGGAATTTGCTGAAGAAAGAAAGATATAAGTATCTCAAGTCTTAGAC 1488
QY 555 SerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnVal 574
DB 1489 CCTAGCCACCGGACATTTCACTGGCTCTTTTGGAGCAGCAGCATGATGAGACAGTT 1548
QY 575 SerTyrProHisLeuValLeuHisGlyValCysMetAlaGly---AspSerIleMet 593
DB 1549 TCCCAACAAACATCGTGTACCTCTACGGCTCTGTGTCGAGATGTAGAAATATATCATG 1608
QY 594 ValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeu 613
DB 1609 GTGAGAGATTTGTGGAGGGGGGCCATTCGATCTCTTCATGCACCGGAAAGGATGCA 1668
QY 614 ValSerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeu 633
DB 1669 CTTTACTACCTCTGGAAGTTTAAAGTTGCAAAACAGCTGGCGCCCTGAGTTACTTG 1728
QY 634 GluAspLysGlyLeuProHisGlyAsnValSerAlaAlaArgLysValLeuAlaArgGlu 653

```

Db 1729 GAAGTAAAGACCTCGTTTCATGGAATGTGTGCACTAAACCTCTCTCTGCGCGGTGAG 1788
Qy 654 GlyIyAspGlyAsn---ProPheIleLeuSerAspProGlyValSerProThr 672
Db 1789 GCATTGACAGTGAATGGCCCGTTTCATAGCTTAGTGAGCTTGGCATCCAGTCTCT 1848
Qy 673 ValLeuSerLeuGluMetLeuThrAspArgIleProTTPValAlaProGluCysLeuGln 692
Db 1849 GTGCTGACGAGCAAGAGTGCATAGAGCAATCCCTGGATCGCTCGCTGAGTGTGTGAA 1908
Qy 693 GluAlaGlnThrLeuCysLeuGluAlaAspIlystTTPGlyPheGlyAlaThrThrTTPGlu 712
Db 1909 GACTCAAGAACCTCAGTGTGGCTCTGCAAGTGGAGCTTTTGGACCACTCTGGGAA 1968
Qy 713 ValPheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaIlystLeuIysPhe 732
Db 1969 ATCCGCTACGATGGCGAGATCCCATCAAGCAAGACCCCTCATTTGAGAAAGAGGTTT 2028
Qy 733 TyrGluAspGlnGlyGlnLeuProAlaLeuIlystTTPThrGluLeuAlaGlyLeuIleThr 752
Db 2029 TATGAAGCCGCTGAGGCGAGTGACACCATCTTCAAGGAGCTAGCTGACCTCATGACT 2088
Qy 753 GlnCysMetAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeu 772
Db 2089 CGTGCATGAATATGATATCCCAACAGAGACCTTTCTCCAGGCCATCATGAGGACAT 2148
Qy 773 AsnGlyLeuIleThrSerAspTyrGluLeuSerAspProThrProGlyIleProSer 792
Db 2149 AACAGCTGGAGGAGCAATCCAGACATGTTTCAGAAAGACCCCA----- 2196
Qy 793 ProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIle 812
Db 2197 -----ATAACAGAGGTGATCCCACTCAT 2220
Qy 813 PheGluGlnArgHisLeuIlystTyrIleSerLeuLeuGlyIysGlyAsnPheGlySerVal 832
Db 2221 TTTGAAAGCGTTTCTTAAAGAGGATTCGTGACTTTGGGAGAGGTCACTTTGGGAAGGTT 2280
Qy 833 GluLeuCysArgTyrAspProGlyAspAsnThrGlyProLeuValAlaValIysGln 852
Db 2281 GAGCTTCGAGATATGATCTTCAGGGAGAACACAGGGGAACAGGTAGCTGTCAAGTCC 2340
Qy 853 LeuGln---HisSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeu 871
Db 2341 CTGAAGCTCGAGAGTGGAGGTAACACACATGCTGATCTGAAGAGGAGATAGATCTTA 2400
Qy 872 LysAlaLeuHisSerAspPheIleValIlystTyrArgGlyValSerTyrGlyProGlyArg 891
Db 2401 CGGAACCTCTACCCAGAGAACATTTGAAAGTATAAAGGAATCTGCATGGAAGACGGGGG 2460
Qy 892 GlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGln 911
Db 2461 AATGGTATCAAGCTCATCATGAGATTTCTGCTTCGGGAAGCCCTAAAGGAATATCTGCCA 2520
Qy 912 ArgHisArgGly---LeuHisThrAspArgLeuLeuPheAlaTTPGlnIleCysIys 930
Db 2521 AAGAATAAGAACAAATACTCAACCTCAACAGCAGCTAAATAATGCCATCCAGATTTGTAG 2580
Qy 931 GlyMetGluTyrLeu 935
Db 2581 GGGATGAGTACTCTCG 2595

```

RESULT 15

```

US-10-131-410-22
; Sequence 22, Application US/10131410
; Publication No. US20030235915A1
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE

```

```

; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST
; FILE OF INVENTION: TUMORS
; FILE REFERENCE: SCH-1763
; CURRENT APPLICATION NUMBER: US/10/131,410
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 09/646,673
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: PCT/DE99/00908
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 2327
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-410-22

Alignment Scores:
Pred. No.: 1,64e-66 Length: 2327
Score: 746.00 Matches: 151
Percent Similarity: 64.06% Conservative: 54
Best Local Similarity: 47.19% Mismatches: 97
Query Match: 12.73% Indels: 18
DB: 12 Gaps: 3

US-09-397-967A-16 (1-1099) x US-10-131-410-22 (1-2327)
Qy 769 LeuArgAspLeuAsnGlyLeuIleThrSerAspTyrGluLeuLeuSerAspProThrPro 788
Db 8 GTCCGAGACATTAATAGCTTGAAGAGCAGATCCAGATATTGTTTCAGAAAAAACCACCA 67
Qy 789 GlyIleProSerProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGln 808
Db 68 GCAACTGAA-----GTG 79
Qy 809 AspProAlaIlePheGluGlnArgHisLeuIlystTyrIleSerLeuLeuGlyIysGlyAsn 828
Db 80 GACCCACACATTTTGAAGAGCGCTTCTTAAAGAGAGTCCGTGACTTTGGGAGAGGCCAC 139
Qy 829 PheGlySerValGluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuVal 848
Db 140 TTTGGGAAGGTTGAGCTCTGAGGTATGACCCCGAAGGGCCATACAGGGGAGCAGGTG 199
Qy 849 AlaValIysGlnLeuGln---HisSerValProAspGlnGlnArgAspPheGlnArgGlu 867
Db 200 GCTGTTAAATCTCTGAAGCGCTGAGAGTGAACCCACATAGCTGATCTGAAAAAGGAA 259
Qy 868 IleGlnIleLeuIysAlaLeuHisSerAspPheIleValIlystTyrArgGlyValSerTyr 887
Db 260 ATCGAGATCTTAAGGAACCTCTATCATGAGAACATTTGAAGTACAAAGGAATCTGCACA 319
Qy 888 GlyProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArg 907
Db 320 GAAGACGGAGGAAATGCTATTAAAGCTCATCATGGAATTTCTGCTTCGGGAAGCCTTAAG 379
Qy 908 AspLeuLeuGlnArgHisArgGly---LeuHisThrAspArgLeuLeuPheAlaTTP 925
Db 380 GAATATCTTCAAGAAATAAGAACAAATAAACCCTCAACAGCAGCTTAAATAATATGCGGTT 439
Qy 927 GlnIleCysIysGlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAla 946
Db 440 CAGATTTGTAAGGGGATGGACTATTGGGTCTTCGGCAATACGCTTCCCGGACTTTGGCA 499
Qy 947 AlaArgAsnIleLeuValGluSerGluAlaHisValIlystIleAlaAspPheGlyLeuAla 966
Db 500 GCAAGAAATGCTCTTGTGAGAGTGAACCAAGTGAATAATGGAGACTTCGGTTTAACC 559
Qy 967 LysLeuLeuProLeuGlyIysAspTyrTyrValValArgGluProGlyGlnSerProIle 986
Db 560 AAAGCAATTGAAACCGATGAAGAGTATTACCCGCTCAAGGATGACCCGGGACACGCCCTGTG 619
Qy 987 PheTyrTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTTP 1006

```

Thu Feb 5 11:06:35 2004

620	TTTTGGTATGCTCCAGAAATGTTTAAATGCAATCTAAATTTTATATTCCTCTCAGCTCTGG	579
Db		
1007	SeqPheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSer	1026
QY	TT	
Db		
680	TCCTTTTGGAGTCACTCTGAGTAGAGTGCTGACTTACTGTGATTTCAGATTCTAGTCCCATG	739
Db		
1027	AlaGluPheLeuArgMetMetGlyProGluAArgGluGlyProProLeuCysArgLeuLeu	1046
QY	TT	
Db		
740	GCTTTTGTCTCGAAATGATAGGCCCAACCCATGGCCAGATGACGTACACAGACTGTGTG	799
Db		
1047	GluLeuLeuAlaGluGlyAArgArgLeuProProProThrCysProThrGluValGln	1066
QY	TT	
Db		
800	AATACGTTAAAGAGGAAACGCCTCGCGTCCCACTAACTGTCCAGATGAGGTTTAT	859
Db		
1067	GluLeuMetGlnLeuCysTyrTrpAlaProGluLupProHisAspArgProAlaPheAlaThrLeu	1086
QY	TT	
Db		
860	CAACTATAGGAAATCTCGGAATTCACACCATCCAACTCGGACAGACTTTAGAACCTT	919
Db		

Search completed: February 4, 2004, 11:29:00
Job time : 1207.04 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2004, 11:29:09 ; Search time 142 Seconds
(without alignments)
3416.057 Million cell updates/sec

Title: US-09-397-967A-16
Perfect score: 1099
Sequence: 1 MAPSEETPLIPRSCSLSS.....RPAFATLSPQLDPLWRGPG 1099

Scoring table: OLIGO

 Xgapop 60.0 , Xgapext 60.0
 Xgapop 60.0 , Xgapext 60.0
 Xgapop 6.0 , Xgapext 7.0
 Xgapop 6.0 , Xgapext 7.0

Searched: 569978 seqs, 220691566 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1135299

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09397967/runat_03022004_175717_28949/app_query.fasta.1.1287
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=olip2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967@cgn 1 1 71 @runat_03022004_175717_28949 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	180	16.4	4016	5	PCT-US95-08354A-1
2	62	5.6	3620	4	US-09-016-434-1055
3	62	5.6	3807	1	US-08-357-598-1
4	62	5.6	3807	2	US-09-003-289-1
5	62	5.6	3807	5	PCT-US95-16435-1
6	16	1.5	3435	3	US-09-046-158A-21
7	16	1.5	3495	1	US-08-446-038B-2
8	16	1.5	3495	1	US-08-446-010B-2
9	16	1.5	3495	1	US-08-805-445-2
10	16	1.5	3495	2	US-08-064-067D-2
11	16	1.5	3495	2	US-09-066-208-2
12	16	1.5	3495	4	US-08-980-080-3
					Sequence 1, Appli
					Sequence 1055, Ap
					Sequence 1, Appli
					Sequence 1, Appli
					Sequence 21, Appli
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 3, Appli

13	1.5	3629	1	US-08-097-997A-8	Sequence 8, Appli
14	1.5	3629	3	US-08-665-574C-8	Sequence 8, Appli
15	1.5	3629	3	US-08-946-994-8	Sequence 8, Appli
16	1.5	4078	4	US-09-016-434-297	Sequence 297, App
17	1.5	4482	2	US-08-567-508C-1	Sequence 1, Appli
18	1.5	4482	3	US-09-196-480-1	Sequence 1, Appli
19	1.5	5117	4	US-09-972-800A-15	Sequence 15, Appli
20	1.2	41	1	US-08-481-003-9	Sequence 9, Appli
21	1.2	41	3	US-08-485-598-9	Sequence 9, Appli
22	1.2	45	1	US-08-481-003-8	Sequence 8, Appli
23	1.2	45	3	US-08-485-598-8	Sequence 8, Appli
24	1.1	151	1	US-08-222-616-11	Sequence 11, Appli
25	1.1	151	4	US-08-446-648-11	Sequence 11, Appli
26	1.1	151	5	PCT-US95-04228-11	Sequence 11, Appli
27	1.1	522	4	US-09-698-505A-29	Sequence 29, Appli
28	1.1	738	2	US-08-604-989A-8	Sequence 8, Appli
29	1.1	927	4	US-10-004-542-3	Sequence 3, Appli
30	1.1	1398	2	US-08-604-989A-9	Sequence 9, Appli
31	1.1	1521	2	US-08-604-989A-10	Sequence 10, Appli
32	1.1	1713	4	US-09-741-154-1	Sequence 1, Appli
33	1.1	1942	2	US-08-604-989A-11	Sequence 11, Appli
34	1.1	1987	2	US-08-876-882-1	Sequence 1, Appli
35	1.1	1987	4	US-09-315-928-1	Sequence 1, Appli
36	1.1	2000	4	US-08-426-509A-1	Sequence 1, Appli
37	1.1	2000	4	US-08-232-545-1	Sequence 1, Appli
38	1.1	2000	5	PCT-US95-05008-1	Sequence 1, Appli
39	1.1	2820	1	US-08-162-809-5	Sequence 5, Appli
40	1.1	2829	4	US-10-004-542-1	Sequence 1, Appli
41	1.1	2962	2	US-08-449-645A-10	Sequence 10, Appli
42	1.1	2962	2	US-08-702-367A-10	Sequence 10, Appli
43	1.1	2962	5	PCT-US95-04681-10	Sequence 10, Appli
44	1.1	2982	1	US-08-348-143-2	Sequence 2, Appli
45	1.1	2982	1	US-08-571-785-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
PCT-US95-08354A-1
; Sequence 1, Application PC/TUS9508354A
; GENERAL INFORMATION:
; APPLICANT: Temple University - Of The
; APPLICANT: Commonwealth System of Higher Education
; TITLE OF INVENTION: JAK3 PROTEIN TYROSINE
; TITLE OF INVENTION: KINASE AND DNA ENCODING THE SAME
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna
; ADDRESSEE: & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08354A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/272,368
; FILING DATE: 8 July 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-203 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4016 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: single stranded
 TOPOLOGY: linear
 PCT-US95-08354A-1

Alignment Scores:
 Pred. No.: 9, 6e-154 Length: 4016
 Score: 180.00 Matches: 529
 Percent Similarity: 97.24% Conservative: 0
 Best Local Similarity: 97.24% Mismatches: 9
 Query Match: 16.38% Indels: 15
 Gaps: 0

US-09-397-967A-16 (1-1099) x PCT-US95-08354A-1 (1-4016)

QY 535 ArgArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAsp 554
 DB 2257 AGGCGGAGGTCTGGATGGTGGACACATGACTCGAAGTCTCTGAGGTCTGGAC 2316
 QY 555 SerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnVal 574
 DB 2317 TCCAGACATCGAAGTCTGGATGCTTTCTGGAAGCGCAAGCTTGATGAGCCAAAGTA 2376
 QY 575 SerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetVal 594
 DB 2377 TCTACCCGACCTGGTGTACTGACGGCGTCTGCATGGTGGAGACGATCGGTG 2436
 QY 595 GlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuVal 614
 DB 2437 CAGGAATTTGTGTATCTAGAGCAATTTGACATGCTGCGCAAGCGTGGCCACCTGGTG 2496
 QY 615 SerAlaSerTriPlysLeuGlnValThrLysGlnLeuAlaIleAlaLeuAsnTyrLeuGlu 634
 DB 2497 TCGCCAGCTGGAACTGACAGGTGACCAAGCAGCTGGCATATGCCCTTAATCTTGGAG 2556
 QY 635 AspLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGly 654
 DB 2557 GACAAAGGCTCTCTCAGCGCAACGCTCTCAGCAGCAAGGTGCTCGTGGCTCGTAGGGG 2616
 QY 655 GlyAspGlyAsnProPheIleLysLeuSerAspProGlyValSerProThrValLeu 674
 DB 2617 GGTGATGGGAATCCACCTTTCAATAGCTGAGTGAATCTGGTGTGCTGCTGCTGCTGCTG 2676
 QY 675 SerLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAla 694
 DB 2677 AGCTGGAATGCTCAGCAGCAATATCCCTGGTGGTGGCGCCCGAATGCTCTCCAGGAGCT 2736
 QY 695 GlnThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrTrpGluValPhe 714
 DB 2737 CAGACATCTGCTTGGAGGTGACAAAGTGGGCTTTGGAGCCACCGTGGAGGTGT- 2795
 QY 715 GlnArg-GlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGln 734
 DB 2796 CAGCGGGGACCCCGCCACATCACTGCTGGAGCCCGCCAAAGGTGAGTTCATGA 2855
 QY 734 uAspGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCys 754
 DB 2856 GGACCGGACAGCTGCGCGCTCTCAATGGACAGAACTGGCGGAGCTTATCACACAGT 2915
 QY 754 sMetAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGly 774
 DB 2916 CATGGCGTATGATCTGGCGCGCGCTCTCTCGAGCTATCTCTCAGAGACCTCAACGG 2975
 QY 774 yLeuIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProAr 794
 DB 2976 CCTCATTCATCAGATTACAGCTCTCTCA-GACCCACACCTGGCATCCCGAGTCTCTCG 3034
 QY 794 gAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGln 814
 DB 3035 AGATGAGCTGTGCG--T--GGCGCCAGCTCTATGCTGCGCAGGAGCCCGCCATATC-GA 3089

QY 814 uGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeu 834
 DB 3090 GGAGAGACACCTTAAGTACATCTCTTCTGGCAAGGCGCAACTTTGGCAGCGTGAGCT 3149
 QY 834 uCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGly 854
 DB 3150 GTGCCGCTATGACCCCT--TG-GACAATACGGGAGCCCTGTGGCAGTGAACACAGCTACA 3206
 QY 854 nHisSerVal-ProAspGlnGlnArgAspPheGlnArgGluIleGlnLeuLysAlaL 874
 DB 3207 GCACAGCGG-GCCAGACCCAGAGAGGACTTCCACGGGAGATTCCAGATCTCTTAAGGCTC 3265
 QY 874 euHisSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerL 894
 DB 3266 TGCACAGCGACTTCAAGTACCGGGAGTCAAGTATGGGCGAGTCCGCCAGGCGC 3325
 QY 894 euArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeu-LeuGlnArgHis 913
 DB 3326 TCGCGTGTGTATGAGTACCTGCCCGAGCGGTGCTGGCAGA-CTTCTGCGAGGCCAT 3384
 QY 914 --ArgGly-LeuHisThrAspArgLeuLeuPheAlaIleGlnIleCysLysGlyMet 932
 DB 3385 CGCGGGCCCTGCACACCGACCGCTACTGCTGCTTGGTGGCAGATCTGCAAGGCGATG 3444
 QY 933 GluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 DB 3445 GAGTACCTGGTGGCGCGCTGCGGTACACCGTGCCTGCTGCGCGCAACATCTTGGTG 3504
 QY 953 GluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGly 972
 DB 3505 GAGAGCGAGGCTCATGTGAAGATCGCGGACTTCCGCTCGCTGCTGCTGCTGCTGCTG 3564
 QY 973 LysAspTyrTyrValValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGlu 992
 DB 3565 AAGGACTACTACTGCTGCTGCGGAGCTTGGCCAAAGCCCATCTTTTGGTATGATGCCGAG 3624
 QY 993 SerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeu 1012
 DB 3625 TCCCTATCTGACAAATCTTCTCCGCAATCTGAGCTGTGGAGTTCGGAGTGGTGTG 3684
 QY 1013 TyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMet 1032
 DB 3685 TAGAGCTCTTCACTACTGCGACAGAGCTGCGAGCCCATCCGCTGAGTCTCTGCGCATG 3744
 QY 1033 MetGlyProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGly 1052
 DB 3745 ATGGGCGCTGAGCGTGNAGGAGCCCGCTCTGCGCGCTCTGCGAGCTGCTGGCAGAGGC 3804
 QY 1053 ArgArgLeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCys 1072
 DB 3805 CGAGCGCTCCACACCTCCGACCTGCCCGAGGTTCCAGGAGCTCATGCGAGCTGTC 3864

RESULT 2

US-09-016-434-1055
 ; Sequence 1055, Application US/09016434
 ; Patent No. 6500938

GENERAL INFORMATION:

APPLICANT: Janice Au-Young
 APPLICANT: Jeffrey J. Seilhamer
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
 TITLE OF INVENTION: PATHWAY GENE EXPRESSION
 NUMBER OF SEQUENCES: 1490
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 STREET: 3174 PORTER DRIVE
 CITY: PALO ALTO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94304
 COMEUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

```

/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/016,434
/ FILING DATE: HEREWITH
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ REFERENCE/DOCKET NUMBER: PA-0002 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 1055:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3620 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GENBANK
/ CLONE: g1039418
/
/ US-09-016-434-1055

```

```

Alignment Scores:
Pred. No.: 1.47e-46 Length: 3620
Score: 62.00 Matches: 96
Percent Similarity: 97.96% Conservative: 0
Best Local Similarity: 97.96% Mismatches: 1
Query Match: 5.64% Indels: 2
DB: 4 Gaps: 0

```

US-09-397-967A-16 (1-1099) x US-09-016-434-1055 (1-3620)

```

Qy 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnLeuValGluSerGluAlaHis 957
Db 2922 CGCGCTGGTGCACCGGACCTGGCGCCGCGAAACATCTCTGTGGAGAGGAGGCACAC 2981
Qy 958 VallysileAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrVa 977
Db 2982 GTCAAGATCGTGACTTCGGCTAGCTAAGCTGCTGGCGTTGAC-AAAGACTACTACGT 3040
Qy 977 lValArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAs 997
Db 3041 GGTCCGCGAGCCAGCCAGAGCCCAATTTCTGTATGCCCCGAAATCCCTCTCGGACAA 3100
Qy 997 nilePheSerArgGlnSerAspValTyrPsePheGlyValValLeuTyrGluLeuPheTh 1017
Db 3101 CATCTTCTCTCGCAGTCAGACGCTCGAGCTTCGGGGTGTCTGTGTACGAGCTCTTCAC 3160
Qy 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034
Db 3161 CTACTCGCAAAAGAGTGCAGCCCTCGCGCGAGTTCTCTCGGATGATGGGA 3212

```

RESULT 3

```

/ US-08-357-598-1
/ Sequence 1, Application US/08357598
/ Patent No. 5705625
/ GENERAL INFORMATION:
/ APPLICANT: Civin, Curt I.
/ APPLICANT: Small, Donald
/ TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
/ NUMBER OF SEQUENCES: 12
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fish & Richardson P.C.
/ STREET: 4225 Executive Square, Suite 1400
/ CITY: La Jolla
/ STATE: CA
/ COUNTRY: USA

```

```

/ ZIP: 92037
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/357,598
/ FILING DATE: 15-DEC-1994
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Haile, Lisa A.
/ REGISTRATION NUMBER: 38,347
/ REFERENCE/DOCKET NUMBER: 07265/033001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 619/678-5070
/ TELEFAX: 619/678-5099
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3807 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA
/
/ US-08-357-598-1

```

```

Alignment Scores:
Pred. No.: 1.54e-46 Length: 3807
Score: 62.00 Matches: 96
Percent Similarity: 97.96% Conservative: 0
Best Local Similarity: 97.96% Mismatches: 1
Query Match: 5.64% Indels: 2
DB: 1 Gaps: 0

```

US-09-397-967A-16 (1-1099) x US-08-357-598-1 (1-3807)

```

Qy 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnLeuValGluSerGluAlaHis 957
Db 2985 CGCGCTGGTGCACCGGACCTGGCGCCGCGAAACATCTCTGTGGAGAGGAGGCACAC 3044
Qy 958 VallysileAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrVa 977
Db 3045 GTCAAGATCGTGACTTCGGCTAGCTAAGCTGCTGGCGTTGAC-AAAGACTACTACGT 3103
Qy 977 lValArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAs 997
Db 3104 GGTCCGCGAGCCAGCCAGAGCCCAATTTCTGTATGCCCCGAAATCCCTCTCGGACAA 3163
Qy 997 nilePheSerArgGlnSerAspValTyrPsePheGlyValValLeuTyrGluLeuPheTh 1017
Db 3164 CATCTTCTCTCGCAGTCAGACGCTTCGGAGCTTCGGGGTGTCTGTGTACGAGCTCTTCAC 3223
Qy 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034
Db 3224 CTACTCGCAAAAGAGTGCAGCCCTCGCGCGAGTTCTCTCGGATGATGGGA 3275

```

RESULT 4

```

/ US-09-003-289-1
/ Sequence 1, Application US/09003289
/ Patent No. 5916792
/ GENERAL INFORMATION:
/ APPLICANT: Civin, Curt I.
/ APPLICANT: Small, Donald
/ TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
/ NUMBER OF SEQUENCES: 12
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fish & Richardson P.C.
/ STREET: 4225 Executive Square, Suite 1400
/ CITY: La Jolla
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 92037
/ COMPUTER READABLE FORM:

```

us-09-397-967a-16.olip2n.rni

Thu Feb 5 11:06:33 2004

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16435
FILING DATE: 15-DEC-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/033W01
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3807 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
PCT-US95-16435-1

Alignment Scores:
Pred. No.: 1.54e-46 Length: 3807
Score: 62.00 Matches: 96
Percent Similarity: 97.96% Conservative: 0
Best Local Similarity: 97.96% Mismatches: 1
Query Match: 5.64% Indels: 2
DB: Gaps: 0

US-09-397-967A-16 (1-1099) x PCT-US95-16435-1 (1-3807)

QY 938 ArgAtgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957
Db 2985 CGCCGCTGCTGACCTTGGCGCTAGCTTAAAGCTCTGCGCTTGCAC-AAAGACTACTACGT 3103

QY 958 ValLysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrVa 977
Db 3045 GTCAAGATCGTGCATTCGGCGCTAGCTTAAAGCTCTGCGCTTGCAC-AAAGACTACTACGT 3103

QY 977 lValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLeuSerAspAs 997
Db 3104 GGTCCGCGAGCCAGGCGAGCCGATTTCTGGTATGCCCGCGAATCCCTCTCGGACAA 3163

QY 997 nIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluLeuPheTh 1017
Db 3164 CATCTTCTCTCGCGAGTCAGACGCTTGGAGCTTCGGGGTCTGCTGTACGAGCTCTTAC 3223

QY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034
Db 3224 CTACTGCGACAAAGAGCTGCAGCCCTCGGCGGAGTTCTCGGATGATGGGA 3275

RESULT 6
US-09-046-158A-21
Sequence 21, Application US/09046158A
Patent No. 6187552
GENERAL INFORMATION:
APPLICANT: Roberts, Steven L.
APPLICANT: Kaytes, Paul S.
TITLE OF INVENTION: METHOD FOR IDENTIFYING INHIBITORS OF
TITLE OF INVENTION: JAK2/CYTOKINE RECEPTOR BINDING
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pharmacia & Upjohn Co., Intellectual Property
ADDRESSEE: Legal Services
STREET: 301 Henrietta Street
CITY: Kalamazoo
STATE: MI
COUNTRY: USA
ZIP: 49001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/003,289
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/357,598
FILING DATE: 15-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/033001
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3807 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-003-289-1

Alignment Scores:
Pred. No.: 1.54e-46 Length: 3807
Score: 62.00 Matches: 96
Percent Similarity: 97.96% Conservative: 0
Best Local Similarity: 97.96% Mismatches: 1
Query Match: 5.64% Indels: 2
DB: Gaps: 0

US-09-397-967A-16 (1-1099) x US-09-003-289-1 (1-3807)

QY 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957
Db 2985 CGCCGCTGCTGACCTTGGCGCTAGCTTAAAGCTCTGCGCTTGCAC-AAAGACTACTACGT 3103

QY 958 ValLysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrVa 977
Db 3045 GTCAAGATCGTGCATTCGGCGCTAGCTTAAAGCTCTGCGCTTGCAC-AAAGACTACTACGT 3103

QY 977 lValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLeuSerAspAs 997
Db 3104 GGTCCGCGAGCCAGGCGAGCCGATTTCTGGTATGCCCGCGAATCCCTCTCGGACAA 3163

QY 997 nIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluLeuPheTh 1017
Db 3164 CATCTTCTCTCGCGAGTCAGACGCTTGGAGCTTCGGGGTCTGCTGTACGAGCTCTTAC 3223

QY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034
Db 3224 CTACTGCGACAAAGAGCTGCAGCCCTCGGCGGAGTTCTCGGATGATGGGA 3275

RESULT 5
PCT-US95-16435-1
Sequence 1, Application PC/TUS9516435
GENERAL INFORMATION:
APPLICANT: The Johns Hopkins University School of Medicine
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk


```
; REFERENCE/DOCKET NUMBER: LUD 5244.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-688-9200
; TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3495 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: nucleic acid
US-08-446-010B-2
Alignment Scores:
Pred. No.: 8.97e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
Gaps: 1
US-09-397-967A-16 (1-1099) x US-08-446-010B-2 (1-3495)
QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
DB 2695 TCAGATGTGTGGAGCTTTGGAGTGGTCTCTATACGAACCTTTTCACATAC 2742
RESULT 9
US-08-805-445-2
; Sequence 2, Application US/08805445
; Patent No. 5821069
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemiacki, Andrew;
; APPLICANT: Harpur, Ailsa
; TITLE OF INVENTION: No. 5821069el Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/805,445
; FILING DATE: 25-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,038
; FILING DATE: 19-MAY-1995
; APPLICATION NUMBER: 08/064,067
; FILING DATE: 30-Jun-1993
; APPLICATION NUMBER: PCT/US91/08889
; FILING DATE: 26-No. 5821069-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian PK3594/90
; FILING DATE: 28-No. 5821069-1990
; APPLICATION DATA: Australian 88229/91
; FILING DATE: 27-No. 5821069-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5821069man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5244
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-688-9200
; TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3495 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: nucleic acid
US-08-064-067D-2
Alignment Scores:
Pred. No.: 8.97e-05 Length: 3495
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3495 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: nucleic acid
US-08-805-445-2
Alignment Scores:
Pred. No.: 8.97e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
Gaps: 1
US-09-397-967A-16 (1-1099) x US-08-805-445-2 (1-3495)
QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
DB 2695 TCAGATGTGTGGAGCTTTGGAGTGGTCTCTATACGAACCTTTTCACATAC 2742
RESULT 10
US-08-064-067D-2
; Sequence 2, Application US/08064067D
; Patent No. 5852184
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemiacki, Andrew;
; APPLICANT: Harpur, Ailsa
; TITLE OF INVENTION: No. 5852184el Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/064,067D
; FILING DATE: 30-Jun-1993
; APPLICATION NUMBER: PCT/US91/08889
; FILING DATE: 26-No. 5852184-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian PK3594/90
; FILING DATE: 28-No. 5852184-1990
; APPLICATION DATA: Australian 88229/91
; FILING DATE: 27-No. 5852184-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5852184man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5244
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-688-9200
; TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3495 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: nucleic acid
US-08-064-067D-2
Alignment Scores:
Pred. No.: 8.97e-05 Length: 3495
```

```

Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 2 Gaps: 0

US-09-397-967A-16 (1-1099) x US-08-064-067D-2 (1-3495)

Qy 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
Db 2695 TCAGATGTGGAGCTTTGGAGTGGTCTTATACGACTTTTCACATC 2742

RESULT 11
US-09-066-208-2
; Sequence 2, Application US/09066208
; Patent No. 5910426
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;
; APPLICANT: Harpur, Allsa
; TITLE OF INVENTION: No. 5910426el Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IEM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,208
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/805,445
; FILING DATE: 25-FEB-1997
; APPLICATION NUMBER: US 08/446,038
; FILING DATE: 19-MAY-1995
; APPLICATION NUMBER: 08/064,067
; FILING DATE: 30-Jun-1993
; APPLICATION NUMBER: PCT/US91/08889
; FILING DATE: 26-No. 5910426-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian PK3594/90
; FILING DATE: 28-No. 5910426-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian 88229/91
; FILING DATE: 27-No. 5910426-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5910426man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5244
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-688-9200
; TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3495 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: nucleic acid
US-09-066-208-2

Alignment Scores:
Pred. No.: 8.97e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```

APPLICANT: Ihle, James N.
APPLICANT: Silvennoinen, Ollie
APPLICANT: Witthuhn, Bruce A.
APPLICANT: Quelle, Frederick W.
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal
TITLE OF INVENTION: Transduction
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,997A
FILING DATE: 29-JULY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656.0370000/SLF/GKT
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3629 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 94...3480
US-08-097-997A-8
Alignment Scores:
Pred. No.: 9,29e-05 Length: 3629
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 1 Gaps: 0
US-09-397-967A-16 (1-1099) x US-08-097-997A-8 (1-3629)
QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
Db 3196 TCAGATGTGGAGCTTTGGAGTGTCTATACGAACCTTTTCACATAC 3243
RESULT 14
US-08-665-574C-8
Sequence 8, Application US/08665574C
Patent No. 6136595
GENERAL INFORMATION:
APPLICANT: Ihle, James N.
APPLICANT: Silvennoinen, Ollie
APPLICANT: Witthuhn, Bruce A.
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine
TITLE OF INVENTION: Signal Transduction
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,574C
FILING DATE: 18-JUN-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/282,012
FILING DATE: 29-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/097,997
FILING DATE: 29-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/118,968
FILING DATE: 09-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bugaisky, Lawrence B.
REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET NUMBER: 0656.0370002/SLF/LBB
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3629 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 94...3480
US-08-665-574C-8
Alignment Scores:
Pred. No.: 9,29e-05 Length: 3629
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 3 Gaps: 0
US-09-397-967A-16 (1-1099) x US-08-665-574C-8 (1-3629)
QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
Db 3196 TCAGATGTGGAGCTTTGGAGTGTCTATACGAACCTTTTCACATAC 3243
RESULT 15
US-08-946-994-8
Sequence 8, Application US/08946994
Patent No. 6210654
GENERAL INFORMATION:
APPLICANT: Ihle, James N.
APPLICANT: Silvennoinen, Ollie
APPLICANT: Witthuhn, Bruce A.
APPLICANT: Quelle, Frederick W.
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal
TITLE OF INVENTION: Transduction
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/946,994
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/665,574
FILING DATE: 18-JUN-1996
APPLICATION NUMBER: 08/282,012
FILING DATE: 29-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/097,997
FILING DATE: 29-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/118,968
FILING DATE: 09-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656.0370002/SLF/GKT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3629 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 94..3480
US-08-946-994-8

Alignment Scores:
Pred. No.: 9.29e-05 Length: 3629
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 3 Gaps: 0

US-09-397-967A-16 (1-1099) x US-08-946-994-8 (1-3629)

Cy 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
Db 3196 TCAGATGCTGGACCTTTGGAGTGGTCTATACGAACCTTTTCACATAC 3243

Search completed: February 4, 2004, 17:52:55
Job time : 187 secs

